The Dissertation Committee for Saumya Mohan Certifies that this is the approved version of the following dissertation:

The Quality of Disclosure and Governance and Their Effect on Litigation Risk

Committee:

Laura Starks, Supervisor

Jay Hartzell, Co-Supervisor

Stephen Donald

Paul Tetlock

Sheridan Titman

The Quality of Disclosure and Governance and Their Effect on Litigation Risk

by

Saumya Mohan, B.B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2006

Dedication

To my parents who have been behind me in my every endeavor,
guiding and encouraging me;
my grandparents who by their inspiring personal examples show me daily that nothing is impossible, and that one should always keep growing and learning;
and my dearest husband Avnish, without whose unswerving support, patience and encouragement, this dissertation would never have been possible.

Acknowledgements

My sincerest gratitude goes to my co-advisors, Dr. Jay Hartzell and Dr. Laura Starks, who have been the best advisors a student could have. I also thank various inspiring teachers and academics who have guided me during my years at UT and elsewhere, including Dr. Andres Almazan, Dr. Stephen Donald, Dr. John Griffin, Dr. Paul Tetlock, Dr. Sheridan Titman, Dr. D.V. Ramana, Dr. M. K. Samal, and seminar participants at the University of Texas at Austin, the Eastern Finance Association 2005 and the Financial Management Association 2005.

The Quality of Disclosure and Governance and Their Effect on Litigation Risk

Publication No._____

Saumya Mohan, Ph.D.

The University of Texas at Austin, 2006

Supervisor: Laura Starks

Abstract: This dissertation examines the relationship between three sets of variables: corporate governance and monitoring, the quality of disclosure in annual reports and securities class action litigation. In the first section, I present a game-theoretic model in which shareholders select from ex ante monitoring or ex post litigation mechanisms available to them in order to mitigate the agency problem. Firm characteristics determine the choice of which of these two mechanisms is appropriate for a particular company. I then test predictions from this model and find that firms with poor monitoring are much more likely than those with good monitoring to be sued even after controlling for the common determinants of a lawsuit. The second section of the dissertation relates the quality of disclosure in annual reports to litigation. I use a dataset containing annual reports filed electronically with the SEC in the period 1996-2005. Using two content analysis software programs that analyze the categories of words used in these annual

reports, I find that firms that use more numbers, past and future words, and other informative words are much less likely to be sued, even after controlling for the common determinants of lawsuits. In order to avoid subjectively choosing categories, I use principal components analysis to identify the major components of annual report disclosure. When these components are used as regressors to identify causative factors of lawsuits, one component named '*informativeness*' has significant power to explain subsequent lawsuits. In head-to-head comparisons of the '*informativeness*' principal component with Standard & Poor's Transparency and Disclosure score, my informativeness measure is more effective than the S&P score in predicting the likelihood of a lawsuit. Finally, in cross-sectional tests, I find support for the theory that firms with good boards and managers who are not entrenched have better disclosure practices. Further, monitoring by institutional investors, independent boards and analysts appears to induce better corporate disclosure.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Ex ante monitoring and ex post litigation as substitutes

The primary goal of corporate governance and monitoring is to solve the agency problem that arises when ownership is separated from management, and to assure investors a return on their investment. As Shleifer and Vishny (1997) have noted, "People who sink the capital need to be assured that they get back the return on this capital. Corporate governance mechanisms provide this assurance." With increasing instances in the last few years of corporate wrongdoing and managerial fraud, corporate governance mechanisms have gained more importance and media attention than ever before. Previous research on corporate governance has examined several ways in which managerial incentives can be aligned so as to prevent fraudulent practices, including earnings manipulation, outsider-dominated boards, incentive-based compensation and monitoring by institutional shareholders and blockholders.

Securities Class Action litigation is another means available to small investors to enforce their rights against managers. Securities Class Action lawsuits are initiated by one or more shareholders representing a larger group (the "class") of shareholders who suffered losses caused by managerial actions. These lawsuits generally involve allegations that managers have engaged in some type of malfeasance such as disseminating false or misleading information, insider trading or earnings manipulation with the result that shareholders have suffered erosion in the market value of their shares.

With the Private Securities Litigation Reform Act of 1995, securities class actions have played an increasing role in disciplining managers. Both the number of cases and

the amounts of settlements have increased dramatically in the last 7 or 8 years.[1] Not only do the companies pay a price on losing a lawsuit, but the management also faces stringent consequences. Niehaus and Roth (1999) and Strahan (1998) both show that CEO turnover increases after securities class action suits, and this effect is stronger after a shareholder victory. Fich and Shivdasani (2004) show that fraud and subsequent lawsuits result in reputational penalties for directors as well. Subsequent to the revelation of fraud, directors on the boards of the fraudulent firms have lost their directorships not only in those firms, but other firms as well.

Since Securities Class Action (SCA) suits impose strict consequences upon discovery of manipulation and fraud, it is clear that securities litigation is one of many corporate governance mechanisms that monitor and discipline managers. In this paper, I model and test the prediction that Securities Class Actions are an ex post substitute to ex ante monitoring as a means of keeping management in check and preventing fraudulent practices. Hence SCAs should come into force especially in those firms in which the other controls are absent or ineffective – namely, in firms that have poor monitoring mechanisms in place. This is the unique contribution of this paper to the literature.

The theoretical model in Chapter 2 characterizes the decision of shareholders who have two mechanisms available to them: ex ante monitoring and ex post litigation. If shareholders choose a high level of ex ante monitoring, it reduces managerial opportunity to commit fraud. Alternatively, the threat of ex post litigation works by discouraging managerial malfeasance. In equilibrium, these two mechanisms are found to be substitutes. The optimal choice for a firm's shareholders depends on i) the cost of

[1] Large settlements in the past few years include Cendant Corporation ($2.85 billion), Waste Management Corp. ($220 million), 3Com ($259 million), Rite-Aid ($200 million), Informix ($159 million) and Prison Realty ($110 million).

monitoring for that particular firm and ii) the ease of ex ante monitoring as opposed to the challenge of proving a charge of fraud in the courts. Comparative statics from the theoretical model give rise to the testable proposition that monitoring by institutional blockholders or boards can be a substitute to litigation. Monitoring of managers is easier in firms with large institutional blockholders or outsider-dominated boards and hence these firms depend less on litigation. Additional comparative statics of the model identify other factors that are likely to determine which of the two is chosen by a company. First, shareholders of firms with entrenched managers find it harder to monitor, and therefore depend heavily on litigation. Second, shareholders of extremely large and complex firms find it harder to monitor their managers and hence depend more on litigation.

To test the predictions of this model, I first examine whether litigation is a substitute to monitoring by three different parties frequently known to act as monitors, independent directors on the board of a company, institutional blockholders (institutional investors holding 5% or more of the shares of a company), and analysts. The results suggest that firms tend to choose one of either litigation or internal monitoring (by boards or institutional investors), consistent with their being substitutes. This is not the case with analyst monitoring, as firms that have a large analyst following are not significantly less likely to face a securities class action lawsuit, after controlling for size and other determinants of lawsuits. However, firms with one or more large institutional blockholders, or firms with a high proportion of independent directors on their boards are less likely to face a lawsuit. I also examine the choice between monitoring and litigation for different types of firms. Consistent with the hypothesis that litigation is preferred to monitoring for companies with entrenched managers, I find that firms with high CEO

compensation and incentive pay (both absolute and adjusted), as well as firms with lower dividend yield are more likely to face class action lawsuits.

A key decision for shareholders who choose to rely on litigation instead of monitoring as a control is the decision of when to sue. One event that could arouse shareholders' suspicions of managerial wrongdoing is an earnings restatement. Well-documented evidence exists that managers sometimes falsely represent earnings to be higher than they are, resulting in inflated share prices. On subsequent discovery of the manipulation, earnings are restated causing a steep drop in stock price.[2]

However, restatements can also be caused by a change in accounting practices or genuine errors that are identified and corrected. Ex post, it is difficult to accurately distinguish benign restatements from those due to intentional mis-statements. Hence, upon observing a restatement in my model, a shareholder gives companies with good ex ante monitoring mechanisms the benefit of the doubt, and is more likely to sue companies with poor existing monitoring mechanisms. If a strongly monitored company restates its financials, shareholders are more likely to attribute it to a genuine error. This is consistent with the model in which, if shareholders decide to monitor ex ante, they drastically reduce the manager's opportunity of committing fraud and the probability of then observing a restatement. Of course, strong controls are costlier to maintain; if shareholders instead choose to 'wait and see' without monitoring, they are more likely to

[2] An abundance of literature exists on the negative stock price reactions to (and the consequent losses to shareholders arising from) earnings restatements. Recent studies include Owens, Lin and Rogers (2002), Wu (2002), GAO (2002), Moriarty and Livingston (2001) and Palmrose et al. (2002). As pointed out by Palmrose, Richardson and Scholz (2001), a restatement can trigger an SEC investigation, lead to replacement of top executives, and result in the firm being significantly penalized by investors.

sue if they see a restatement or a sharp price drop, because this event is more likely to be a sign of mismanagement for a company with poor corporate governance systems.

The importance that shareholders give to accurate earnings statements highlights the importance of accurate and timely managerial communication with shareholders. The next section describes an analysis of the quality of disclosure in managerial communications with shareholders.

Using disclosure quality to predict litigation

Corporate disclosure plays an integral role in alleviating the problem of information asymmetry between managers and shareholders in a firm. It is the means by which managers communicate information both relating to past events and future growth prospects of a firm. As such, this information flow forms the lifeblood of any capital market. Developing a precise metric for the quality of disclosure by a company would be beneficial because it sets a standard for measuring the openness and accuracy of communications by the management of a company. In this section, I outline a study that analyzes the 10-K filings of firms in order to develop a measure of the informativeness of the communications by a firm.

My measure for the quality of disclosure differs from previous studies that have relied upon an analysis of the accounting information in these documents. Instead, I analyze the types of words used in these documents. Several studies in the psychology area have documented that the usage of certain types of words can be related to attributes like transparency and open communication. Along these lines, I use content analysis software programs to assess the quality of disclosure in these filings. I also relate this measure to the likelihood of a securities class action lawsuit, arguing that these lawsuits

exemplify poor disclosure practices that are rejected by investors. These software programs have built-in dictionaries and wordlists that are used to classify the words in documents they read. Using the results from content analysis software programs, I find that firms that disclose less information about past and future corporate events, and present less numerical data are more likely to be subsequently sued. However, these firms do not have unusual amounts of positive or negative words, or optimistic or pessimistic words. I also use principal components analysis of the more than 200 word categories in the content analysis packages in order to determine the combinations of word types that are most likely to trigger a lawsuit. The result of this analysis is an '*informativeness*' component that is found to be inversely related to the risk of a lawsuit (and hence to poor disclosure practices). I also compare the efficacy of this informativeness component in predicting lawsuits against another measure of disclosure, the S&P Transparency and Disclosure rank. In head-to-head tests, I find that the informativeness component is a better measure of securities litigation risk compared to the S&P rankings based on annual reports and on the entire spectrum of corporate transparency.

Finally, I examine the determinants of good disclosure as measured by the informativeness of 10-K filings of a company. I examine whether good governance practices influence good disclosure practices, by regressing the informativeness component on governance measures including board size and structure, the number of antitakeover provisions and executive compensation measures among others. Results from this test show that firms with strong governance as measured by these structures have better disclosure. In other words, firms with independent boards and less entrenched

CEOs have better disclosure practices. I do not find strong evidence that the size and composition of the CEO pay package is related to the quality of disclosure by a firm.

I also examine the relationship between informative disclosure practices and monitoring by institutional investors and analysts. I try to resolve the causality in this relationship which is ambiguous. Good disclosure practices might attract institutional investors and analysts to a stock; alternatively, when a large number of institutions and analysts have a stake in a company, they may enforce better disclosure. I address this issue of causality using an instrumental variable approach, and instrument for institutional holdings and analyst coverage using membership in the S&P500 index and other indices, and share turnover. Using this approach, I find that firms with a higher proportion of shares held by institutional investors, or firms with at least one institutional blockholder tend to have a better quality of disclosure in their 10-K filings. I find limited evidence to indicate that firms that are followed by more analysts also tend to have better disclosure practices.

This work draws upon and links a number of threads in the accounting manipulation, securities law and corporate governance literature. Chapter 2 describes the existing literature in the areas of securities litigation, fraud, disclosure and governance and lists the contributions of this dissertation to the literature. Chapter 3 outlines a theoretical model which models shareholders' choice of ex ante and ex post measures of monitoring managers' actions. Chapter 4 details the empirical analysis that tests the predictions of the model. Chapter 5 describes the use of disclosure quality measures in order to predict securities litigation. Chapter 6 examines whether common corporate

governance and monitoring mechanisms are determinants of disclosure quality and the last section concludes.

Chapter 2: Literature Review

Governance, Disclosure and Litigation Risk

Securities Class Action Litigation

Introduction to securities class action litigation

The literature on securities class action litigation can broadly be classified into descriptive literature and analytical studies. Several descriptive papers analyze the types of securities lawsuits, the main allegations, and the time trends of settlements[3] and dismissals. Bajaj, Mazumdar and Sarin (2000) and others document the increasing time trend in securities class action litigation. Among the analytical studies, a set of papers examines stock market reactions to litigation announcements. Gande and Lewis (2006), a recent paper that looks at this issue, confirms previously reported negative stock price reactions to shareholder initiated class action lawsuits, on the announcement date, but contends that filing date effects understate the magnitude of shareholder losses because shareholders partially anticipate these lawsuits and capitalize part of these losses prior to a lawsuit filing date. Another group of papers examines the efficacy of legal reform in preventing non-meritorious lawsuits from coming to trial and issuing valid judgments to meritorious lawsuits.

[3] Cornerstone Research publishes securities class action litigation summaries annually, describing the kinds of lawsuits, main allegations and settlements. NERA also publishes a similar review. Other descriptive papers are Carleton, Weisbach and Weiss (1996) and Bajaj, Mazumdar and Sarin (2000).

Several papers also document that the Public Securities Litigation Reform Act (PSLRA) increased the significance of certain merit-related factors, such as the presence of an accounting restatement, in determining the incidence and outcomes of securities fraud class actions. Choi (2005) takes a closer look at this issue by evaluating whether this documented increase in the importance of merit-related factors post-PSLRA has reduced the incidence of both nuisance litigation as well as some meritorious claims. The essay provides evidence that the pre-PSLRA non-nuisance claims lacking obvious "hard evidence" indicative of fraud (an accounting restatement or SEC action) face both a lower probability of suit in the post-PSLRA period and a greater likelihood of receiving a dismissal or low-value settlement in the post-PSLRA period.

Accounting restatements and litigation

Richardson, Tuna and Wu (2002) find that restating firms have higher ex ante financing needs, large total accruals, higher P/E and M/B ratios. Agrawal and Chadha (2005) do not find a strong relationship between the probability of restatement and different measures of board and auditor independence. A few recent papers have examined the connection between restatements and litigation. Palmrose and Scholz (2004) ask the same questions as I do in evaluating why some restatements are more likely to be followed by litigation than others. However, they do not take into consideration agency problems or other corporate governance controls. Their hypothesis deals with features of the restatement itself (pervasiveness, core elements or non-core elements of the financial statements that are restated) that determines whether a restatement results in litigation or not. One caveat with this approach is that it is often hard to distinguish exact features of a restatement from a Lexis-Nexis news report.

Though most reports carry the date of the announcement, the period for which financial statements are restated and the effect on net income or earnings per share, it often requires plenty of skill and an in-depth knowledge of accounting for an investor to look beyond these and assess how pervasive the restatement is.

Agency problems and litigation

In the litigation literature, Niehaus and Roth (1999) look at various measures in order to decide whether securities class actions have merit or are just instigated by self-serving lawyers. They conclude that in several of the litigated firms, managers had incentives for delayed disclosure of negative earnings news and that securities class actions on average do have justification. Strahan (1998) tests (and concludes) that firms prone to agency problems are more likely to face securities class actions. However Strahan does not control for earnings manipulation which is the link between agency problems and litigation. Agency problems make firms more vulnerable to earnings manipulation by self-serving managers, and firms that manipulate earnings are more likely to be sued.

Other factors affecting litigation

Jones and Weingram (1998) analyze the effects of five factors on the likelihood of a company being subject to a lawsuit under Sec. 10b-5 (securities lawsuit) after a major single-day stock price decline: insider trading, seasoned equity offerings, accounting restatements, SEC enforcement actions, and fall-triggering announcements. They investigate the incremental impact of each factor after controlling for stock market characteristics that influence litigation risk. They find that insider sales do not increase

litigation risk. Issuing equity also does not have a significant impact on a corporation's likelihood of being a target of this type of litigation. However, the type of announcement that accompanies the single-day fall influences the likelihood of litigation. Firms that correct previous accounting statements and those that are subject to SEC enforcement actions are substantially more likely to be sued than other firms.

Consequences of litigation

The financial consequences of litigation can be enormous, as can be seen from the experiences of prominent companies in the news lately. Some examples of securities lawsuits with extremely large settlements are Bank of America (490 million), 3Com (259 million), Waste Management (457 million), Qwest Corporation (250 million), MCI/WorldCom (750 million) and Cendant Corporation (2.85 billion). The non-monetary consequences of securities litigation have also been substantial. Both Niehaus and Roth (1999) and Strahan (1998) find increased managerial turnover after a lawsuit. Fich and Shivdasani (2004) report that outside directors tend to lose their directorships following a lawsuit. These and other consequences of fraud (whether followed by litigation or not) are discussed in the section entitled "Consequences of accounting manipulation and fraud".

Accounting Manipulation and Fraud

Managerial incentives to commit accounting manipulation and fraud

One of the seminal papers that examined factors predicting earnings management was Dechow, Sloan and Sweeney (1995). This paper uses discretionary accruals-based models to successfully predict earnings management and SEC fraud. Dechow, Sloan and

Sweeney (1996) examine the causes and consequences of earnings manipulation by analyzing a sample of firms that were subject to SEC enforcement actions. They find that firms manipulating earnings are: (i) more likely to have boards of directors dominated by management; (ii) more likely to have a CEO who simultaneously serves as Chairman of the Board; (iii) more likely to have a CEO who is also the firm's founder; (iv) less likely to have an audit committee; and (v) less likely to have an outside blockholder. They also find that an important motivation for earnings manipulation is the desire to attract external financing at low cost.

Beneish (1999) and Kedia (2003) argue that stock options generate incentives to manipulate earnings.[4] Other papers that associate high or incentive-based CEO pay with earnings management and manipulation are Peng and Roell (2004), Richardson, Tuna and Wu (2002) and Wu (2002). Summers and Sweeney (1998) also concludes that managers who perpetrate fraud tend to sell their stakes in the company and benefit from the stock price manipulation.

Consequences of accounting manipulation and fraud

Desai, Hogan and Wilkins (2006) examine management turnover and the subsequent re-hiring of displaced managers at firms announcing earnings restatements during 1997 or 1998. In contrast to prior research[5] which does not find increased turnover following GAAP violations or revelation of corporate fraud, they find that 60% of restating firms experience a turnover in at least one top manager within 24 months of the restatement compared to only 35% among age-, size- and industry-matched firms.

[4] This conclusion finds support in many other studies including Kedia (2003) and Denis, Hanouna and Sarin (2005).
[5] Beneish (1999) and Agrawal, Jaffe and Karpoff (1999)

Moreover, 85% of the displaced managers of restatement firms are unable to secure comparable employment afterwards. This conclusion is supported by Livingston (1996), who finds that after controlling for firm performance and financial distress, top managers and financial officers are more likely to be dismissed in the years following misleading reporting than in other years. For top executives, SEC enforcement action is associated with a higher frequency of turnover.

Cox and Thomas (2003) examine the relationship between SEC enforcement actions for fraud and private securities litigation by shareholders. Their findings include the following: i) private suits with parallel SEC actions settle for significantly more than private suits without such proceedings; ii) SEC enforcement actions target significantly smaller companies than private actions alone; iii) private cases with parallel SEC actions take substantially less time to settle than other private cases; and iv) private cases with parallel SEC actions have significantly longer class periods than other private actions. In addition, they have the important result that financial distress is the single most important determinant of SEC actions whereas expected losses are not a significant determinant of these actions.

Feroz, Park and Pastena (1991) examine the consequences of an SEC Accounting and Auditing Enforcement Release (AAER) on the firm's managers, auditors and market participants. AAERs are SEC civil actions taken against firms which violate accounting rules, most commonly by overstatements of accounts receivable and inventories resulting from premature revenue recognition and delayed write-offs respectively. Consistent with prior research, they find that the announcement of an SEC AAER results in a significant

negative abnormal return. The AAER is also accompanied by specific SEC censures to the external auditor of the firm regarding particular accounting issues.

The penalties to both managers and directors of fraud and subsequent litigation are stringent. Fich and Shivdasani (2004) have the interesting result that upon revelation of fraud, outside directors are less likely to retain their directorships of both fraud and non-fraud firms. Hence fraud and litigation imposes penalties not only on the company and its management, but also upon its directors.

Disclosure Quality

Measures of disclosure quality

Among the disclosure or "quality of information" proxies suggested in the literature are accounting-based measures and managerial communication-based measures. Francis, Nanda and Olsson (2005) use accruals quality (a measure of how well accruals map into cash flows) as a measure of information quality. This measure is found to be closely correlated with the expansiveness of voluntary information disclosure by the firm. They use this accruals quality proxy to investigate the relationship between information quality and cost of capital, and find that greater voluntary disclosure is associated with lower costs of debt and equity capital.

A second set of papers develop measures of voluntary disclosure based on public statements and announcements by management. Tasker (1998) uses the number of conference calls that provide open access to all investors as a proxy for disclosure quality. Kasznik and Lev (1995) examine whether firms provide early warnings of bad news before previously scheduled earnings announcements. Early warning by firms of

bad earnings news is considered to be good disclosure policy. Kasznik (1999) follows a different approach by examining the accuracy of earnings forecasts by management. Accurate earnings forecasts are treated as a measure of good voluntary disclosure. Waymire (1985), Cox (1985) and Imhoff (1978) are other studies that use management forecasts of earnings as a proxy for voluntary disclosure.

Another set of papers in this field polls outsiders such as analysts to determine their ratings of the company's disclosure policies. The papers that fall into this category include Brown, Finn and Hillegeist (2000) and Lang and Lundholm (1993). Brown et al. (2000) use AIMR score, a rating by the association of analysts as a proxy for voluntary disclosure. However this score is available only until 1995 because AIMR discontinued its ranking. Lang and Lundholm (1993) also use analysts' rating of corporate disclosure informativeness from the 1985-1989 FAF reports. Finally, Easley, Hvidkjaer and O'Hara (2002) adopt a slightly different method by using an index denoted as PIN, or the probability of informed trading as a proxy for information asymmetry. They find that a difference of 10 percentage points in the probability of information-based trading between two stocks leads to a difference in their expected returns of 2.5% per year. They interpret their results as providing strong support for the premise that information affects asset pricing fundamentals.

Standard and Poor's introduced their Transparency and Disclosure (T&D) ranks in 2002.[6] The ranking procedure involved analyzing 98 disclosure items in three categories: ownership structure and investor rights, financial transparency and information disclosure, and board and management structure and process. The S&P

[6] The ranks and ranking procedures are described in the S&P paper by Patel and Dallas (2002).

disclosure rank is based on 35 accounting items, 35 governance structures and 28 ownership structures. The ranks, initially awarded to US companies alone, were later extended to include companies in emerging markets including Latin America and Asia. Both the original study by Patel and Dallas (2002) and other subsequent studies examine the relationship between these S&P ranks and market risk and find a negative relationship between the two. Independently, Cheng, Collins and Huang (2003) concluded that the S&P T&D rankings did provide new information to the markets along various dimensions, and that firms with large differences in disclosure levels in SEC filings over time were received unfavorably by investors.

Finally, Botosan (1997) compares different disclosure scores and constructs one of her own. This score uses information from annual reports of 122 companies and is based on background information of the company, historical results, key non-financial statistics (e.g. number of employees, market share), forecasted information and MD&A. The drawback of this score is that it is only available for a limited number of companies and at one point in time. Thus it does not allow comparisons across time, and allows only limited comparisons of disclosure across firms. Francis, Nanda and Olsson (2005) build on Botosan (1997) by constructing their own score which leaves out MD&A and background information "because disclosure in these categories is substantially constrained by SEC rules" and removing these categories captures voluntary disclosure. They also add a category of other financial information.

Factors affecting disclosure quality

Healy and Palepu (2001) provide a comprehensive review of factors affecting the quality of disclosure as measured by the decision to manage earnings. They discuss both

theoretical and empirical results for six factors that may affect managers' disclosure decisions: other than the threat of litigation discussed in the next section, they also examine capital market transactions, corporate control contests, stock compensation, proprietary costs, and management talent signaling. Theoretical as well as empirical evidence suggests that managers improve their disclosure decisions prior to capital market transactions, in order to avoid hostile takeovers and to reduce the likelihood of undervaluation and explain away poor earnings performance. Managers' compensation plans also provide incentives to engage in voluntary disclosure in order to increase stock liquidity, reveal private information and meet insider trading guidelines. Disclosure decisions are also affected by concerns about revealing firm-specific information in a competitive market and signaling managers' superior ability to anticipate future changes. In contrast, Nagar (1999) characterizes disclosure as a function of the manager's human capital. According to Nagar (1999), a firm's earnings are a function of the assets in place in the firm and the manager's human capital. Disclosure of the earnings by the manager may result in a reassessment of his human capital by the market, possibly affecting his future earnings potential. A manager who is less risk-averse or paid to disclose information is more likely to volunteer disclosure.

Wright (1996) examines the relationship between corporate governance characteristics and the quality of reporting in financial statements. This paper documents significant correlations between various financial reporting quality measures and board composition. The quality of financial reporting is found to be negatively related to the presence of insiders and grey directors on the audit committee. The two financial reporting quality measures evaluated are (1) analysts' evaluations of corporate disclosure

practices and (2) the existence of an SEC Accounting and Auditing Enforcement Release against the firm or its auditors. Moreover, the percentage of insider or grey audit committee members is significantly related to the percentage of stock owned by large shareholders and institutions. This suggests that the role of the audit committee as a monitor of financial reporting practices depends on firm ownership characteristics.

Lang and Lundholm (1996) analyze disclosures specifically for firms that make equity offerings and find that there is a significant increase in disclosure beginning six months before the offering, particularly for the categories of disclosure over which firms have the most discretion. This finding adds to the result in Lang and Lundholm (1993) that firms with more informative disclosures have larger analyst following, less dispersion in analyst forecasts, and less volatility in forecast revisions.

Disclosure quality and litigation

Healy and Palepu (2001) summarize both the theoretical and empirical literature relating disclosure and litigation. Theoretically, it is unclear whether the threat of litigation will lead to improved voluntary disclosure or reduce incentive to disclose forward-looking information. Skinner (1994, 1997) finds that firms with bad earnings news are more likely to voluntarily disclose their earnings performance compared to firms with good news. Francis, Philbrick and Schipper (1994) find that 62% of the firms in their litigation sample were sued over earnings forecasts or pre-emptive earnings disclosures. 87% of their sample of no-litigation firms with comparable stock price declines pre-announced an earnings decline. They conclude that disclosure does not appear to be a deterrent to litigation.

The prior literature has established that a negative earnings report often results in litigation; however there is no consensus on the effect of early disclosure of bad news. Lev (1992) and Skinner (1994) contend that disclosing negative earnings early (prior to regularly scheduled earnings releases) softens the blow and reduces the likelihood of litigation; whereas others such as Francis, Philbrick and Schipper (1994) argue that early disclosure increases the probability of a lawsuit. Field, Lowry and Shu (2005) try to resolve this confusing relationship between disclosure and litigation by asking whether disclosure deters or triggers litigation. By using the simultaneous equations method to control for the endogeneity of disclosure and litigation, they find no evidence to indicate that early disclosure triggers litigation. In fact, they find some evidence that early disclosure of negative news may reduce the possibility of litigation.

Kasznik (1999) in an innovative study examines whether disclosure influences earnings management. The study finds that managers try to reduce the reputational and legal consequences they may face after management earnings forecast errors by using discretionary accruals to manage earnings upward. A consistent finding is that managers use discretionary accruals more in the "post forecast" periods than "pre forecast" periods.

Disclosure quality and corporate governance

Fox (1999) proposes a hitherto unrecognized effect of enforced disclosure policies by arguing that required disclosure helps shareholders enforce managers' fiduciary duties. Other than the direct impact of making managers accountable for their publicly available decisions, enforced disclosure also indirectly affects corporate governance by affecting the market for corporate control, the cost of capital, and monitoring by external sources of finance.

Institutional investors as monitors

Demsetz (1983) and Shleifer and Vishny (1986) were the earliest to propose that the existence of large shareholders will lead to better monitoring of managers. Several papers have since empirically examined whether institutional investors play a monitoring role in mitigating the agency problem between shareholders and managers. Institutional investors influence managerial decisions, and hence play a monitoring role, both actively influencing managerial decisions in various spheres, and passively by selectively buying shares in companies with desirable characteristics. According to Gillan and Starks (2000) institutional investors such as public pension funds began to get actively involved in corporate governance by submitting shareholder proxy proposals as early as 1986. They test if institutional shareholder activism is viewed favorably by investors by examining the relationship between proposal voting patterns and sponsor identity. The test finds that on average, proposals sponsored by institutions get 175% as many votes as those sponsored by individual shareholders.

Other papers have found that institutional shareholders influence managerial pay and business decisions. Hartzell and Starks (2003) find evidence that institutional ownership concentration is positively related to the pay-for-performance sensitivity of executive compensation and negatively related to the level of compensation, suggesting that institutions play a monitoring role. Other evidence suggests that institutions may have influence on business decisions. Bushee and Noe (2000) find that managers are less likely to cut R&D spending when institutional ownership is high, implying that institutional owners do not put pressures on managers to indulge in myopic behavior.

However a high proportion of ownership by institutions that have high portfolio turnover and engage in momentum trading increases the probability that managers will cut R&D spending to reverse an earnings decline.

Clyde (1997) throws some light on the mechanism by which institutional owners influence the company's operations. This paper finds that institutional owners were more likely in the 1980s to police management by using takeovers as the disciplining mechanism. Maug (1998) asks whether a liquid market results in less monitoring by institutional investors because of the ease with which they can dispose of their stock holdings. In fact, the findings are opposite: a liquid market actually enhances monitoring by institutional investors, because it mitigates the free rider problem and makes it easier for large shareholders to hold more shares. Hence liquidity in stock markets is conducive to monitoring by institutional investors. Large institutions like CalPERS involve themselves in promoting good corporate practices through activism. Smith (1996) studies 51 firms targeted by CalPERS and finds that the majority of these adopt the changes proposed by CalPERS.

Institutional investors in securities litigation

Weiss and Beckerman (1995) take a litigation procedure-oriented approach to argue that institutional investors can reduce agency costs in securities lawsuits. They note that institutional investors have a large stake in class actions, since they are the recipient of a large proportion of the payouts. As such, they have an incentive to play an active role in securities litigation, and the legal system should encourage them to become "lead plaintiffs" in securities class action lawsuits. Doing so will benefit the plaintiffs by

increasing the proportion of the damages they receive, and at the same time alleviate concerns about their getting shortchanged by avaricious plaintiff lawyers. This paper was written before the Private Securities Litigation Reform Act (PSLRA) came into effect, putting into place controls for reducing the number of 'strike suits' that come to trial. The PSLRA also made the litigation environment conducive to active institutional participation similar to the suggestions in Weiss and Beckerman.

Johnson (1997) agrees that institutional investors have an important role to play in securities litigation; furthermore, that only institutional investors can fulfill both the roles that securities litigation requires: deterrence of future fraud and adequate compensation for investors who suffered as a result of current managerial fraud. Since these two aims may sometimes be in conflict, institutions are uniquely placed to enhance the deterrence function without undermining the compensation goal.

Compensation and incentives

Ke (2003), for example, finds that CEOs who hold more stock options and stocks are more likely to report longer strings of consecutive earnings increase, especially for firms whose stock prices are historically sensitive to earnings reports. A number of other studies contain similar findings. Johnson, Ryan & Tian (2003) find that executives at fraud firms have significantly larger equity-based compensation and greater financial incentives to commit fraud than do executives at industry- and size-matched control firms. Executives at fraud firms are also shown to benefit from fraud by exercising larger fractions of their vested options than executives in control firms during the fraud years.

My Contribution to the existing literature

In this section I detail the various contributions my dissertation makes to the existing literature on different fronts. As this study spans different research areas including securities class action litigation, corporate disclosure and transparency, governance and monitoring, and fraud, I will list contributions in all these different areas.

The securities class action literature is largely descriptive in nature. An important contribution of my dissertation in the litigation area is to treat litigation as one of the means available to shareholders to enforce their rights. As such, litigation is seen as an "ex post measure" and a substitute for other measures that work "ex ante" to align incentives and keep a check on managerial behavior. Prior literature examining litigation has tended to be in the law area and hence has examined litigation in isolation, and not as one of many complementary structures available to enforce shareholder rights. Similarly, the literature in corporate governance and agency theory has largely ignored litigation as another means available to the same end. I explicitly model this trade-off between ex ante measures like corporate governance and monitoring, and the ex post litigation method, and describe firm characteristics that may determine which of these different methods is suitable for a particular firm.

The second significant contribution of this dissertation is in the area of corporate disclosure and transparency. Though there have been previous attempts to quantify the quality of disclosure in annual reports, these rankings or ratings are based on intensive human analyses of the reports that are hard to replicate. In addition these analyses may have a subjective aspect to them. Further, the existing rankings of disclosure and transparency are available only for specific companies in their sample, and only for

specific years. For example, the AIMR score used by Brown et al. (2000) among others is available only for a limited number of years, ending in 1995 because the rankings were discontinued at that time. Similarly, analysts' ratings of informativeness of corporate disclosure, as studied by Lang and Lundholm (1993) and others, is also available only for the period from 1985 to 1989. The S&P Transparency and Disclosure Rank is the most widespread and detailed of these measures, incorporating information not only from the annual reports but also other filings. However this too is limited in scope as it has only been issued by the S&P for US companies once, in 2001. It is not clear how many years of data were studied in order to determine a company's Transparency and Disclosure rank issued in 2001.

In contrast to all these studies, my disclosure and transparency measure can be calculated for any company, in any year, using a standardized methodology that is programmable and easily replicable. The input data is the 10-K report that is freely available for download on the SEC's website, and the content analysis software programs used in the analysis are also available for free download online. Secondly, and unlike other disclosure measures, I show that my measure of disclosure is strongly and significantly related to the likelihood of a lawsuit. Poor disclosure (based on an analysis of the types of words used in the 10-K document) is likely to lead to a lawsuit. Other measures of disclosure quality are not linked to objective external consequences such as litigation. Securities lawsuits can result in sizeable monetary consequences for the company, reputational penalties for the management and directors, and steep erosion in shareholder value. Hence my measure is not only easily calculated but also economically significant, as 'poor' disclosure has serious negative consequences.

The third significant contribution of this dissertation is analyzing the factors that might predict the quality of disclosure by a firm. There is little literature on the factors that pressure management to improve the quality of their disclosure. I examine whether strong governance and monitoring by institutions and analysts lead to informative disclosure on the part of the management of a company. I find evidence that independent boards, institutional blockholders and analyst following are associated with more informative disclosure levels. This gives us an insight into what steps can be taken in order to improve the informativeness and timeliness of corporate disclosure.

Core (2001) summarized the existing disclosure literature and outlined suggestions for future work which coincidentally shed some light on the scope and contribution of this dissertation in the area of disclosure. The following is a passage from Core (2001)'s suggestions for future work:

> I conjecture that researchers can substantially lower the cost of computing these metrics by importing techniques in natural language processing from fields like computer science, linguistics, and artificial intelligence. An example of a widespread natural language processing technology is the grammar-checking device provided with many word- processing programs. This device provides information on the frequency of use of the passive tense. The passive tense (is) one of a number of linguistic devices for hiding meaning that are examined in the law and linguistics literature. These programs also provide other readability statistics... It seems worthwhile to investigate whether more sophisticated natural language processing technology could be used to replicate ratings by the AIMR and ratings by researchers. If this can be accomplished, it would significantly reduce the cost of creating disclosure quality indices from firm reports and press releases. Natural language processing programs could be also used to create proxies for the ''tone'' of disclosure (Lang and Lundholm, 2001) and proxies for the precision and bias of the information that is conveyed. Healy and Palepu (1993, 2001) emphasize the important idea that managers communicate with investors. Managers use natural language for this communication, and we can advance work in accounting by using research from other fields to find ways to machine-code the precision of this language and any bias contained in it.

I quote this passage because it is a good description of the contribution of this research to the disclosure literature. The 'natural language processing programs', or content analysis programs, are particularly suited for examining the nature of managerial disclosures, and offer the added advantage of being programmable and automated. Thus the disclosure quality index I create can be easily and quickly computed from publicly available filings for thousands of companies and describes different facets of the annual report including optimism or pessimism, proportion of numerical information and forward-looking information etc. The use of this disclosure quality index to predict possible fraud and litigation represents a significant addition to the body of work that examines disclosure, in the finance, accounting and legal fields. The analysis of the determinants of informative disclosure can help investors create conditions in firms that are conducive to good disclosure by management.

Chapter 3

Ex Ante Monitoring or Ex Post Litigation? Theoretical Model

This section describes a theoretical model that characterizes the decision of a shareholder (or group of shareholders) of a company seeking to mitigate the manager-shareholder agency problem. In addressing the potential problem of managerial fraud, shareholders have two mechanisms available to them. One option is to impose ex ante controls so that the manager's opportunity to commit fraud is greatly reduced. They may also choose to let the possibility of an ex post lawsuit act as a deterrent. I analyze shareholders' choices under two different information assumptions. First, to provide a benchmark, I assume that there is perfect information, i.e. that managerial actions including fraud are visible to shareholders. Then, I also examine the more realistic setting of imperfect information where shareholders only receive a signal of the manager's actions, and must decide if they believe that a fraud has been committed. In both settings, shareholders of a company first have to choose between strong and weak monitoring of their manager. If strong monitoring is chosen, the manager is unable to commit fraud as he is given no opportunity to do so. If shareholders choose weak monitoring, they can later decide to litigate if they observe that fraud occurs (in the perfect information setting) or if they have reason to suspect that it has occurred (in the imperfect information setting).

Though fraud itself is unobservable in the real world, we do observe various signals of managerial behavior, including the one I focus on here, restatements of the firm's financial statements. A restatement is a noisy signal of fraud, since a restatement

could result either from fraud having been committed or from a genuine accounting error. Thus, when a shareholder observes a restatement, she has to deduce whether it is attributable to a fraud or an error, and accordingly decide whether to sue. In my model, the shareholder bases this decision on the type of monitoring in the firm (weak or strong). This is because a strongly monitored firm has structures in place that reduce a manager's opportunity to commit fraud. Thus, the probability of a strongly monitored firm committing fraud is lower; and the restatement is more likely to come from an accounting error than from fraud. However in the case of a weakly monitored firm, few structures exist to prevent the manager from committing fraud. Therefore, if such a firm restates its financial statements, the probability of fraud is higher and a shareholder is more likely to sue.

This model assumes that the legal system establishes the manager's guilt or innocence beyond all doubt, both under the perfect information and imperfect information settings.[7] A manager guilty of fraud is punished with a monetary penalty and a wrongly accused manager is cleared and released. However the shareholder bears a cost of suing a manager suspected of fraud. This cost can be thought of as the money spent on procuring lawyers and the time and energy involved in filing a claim; and is incurred whether the suit is successful (upon which the shareholder earns damages - a cash award) or not. This cost attached to suing prevents frivolous lawsuits, making it more likely that only legitimate cases come to trial. The primary prediction from the model is that shareholders either choose to have ex ante strong monitoring, or to have weak monitoring and resort to ex post litigation. These two mechanisms operate as substitutes. Further, the

[7] However this is not essential for the results of the model to go through. As long as the payoff to a manager from committing fraud and being sued is lower than the payoff from not committing fraud, the results will still hold.

assumption of heterogeneous monitoring costs leads to other interesting predictions about whether monitoring or litigation is the optimal choice for a firm based on firm-specific characteristics.

Baseline perfect information model

For simplicity, let us assume that there is only one shareholder who owns all the shares of the company and makes all the decisions. In this first simple model I assume that there exist two possible levels of monitoring: strong or perfect monitoring (M_H) and weak or no monitoring (M_L).[8] The sequence of actions is as follows. The shareholder first can select weak monitoring (which does not cost anything) or strong monitoring, which comes at a cost c that varies by firm. The shareholder's choice of strong or weak monitoring determines the manager's opportunity to commit fraudulent action. Selection of strong monitoring M_H gives the manager no opportunity to commit fraud. (This simplistic model assumes perfect monitoring can be ensured by selecting M_H). If weak monitoring is chosen, the manager has the opportunity to commit fraud, which gives him a private benefit of ξ. 'F' and 'S' are the decision variables that describe the actions of the manager and shareholders, respectively. The variable 'F' is assigned a value of '1' if the manager commits fraud and '0' otherwise. The shareholder can then decide whether to sue or not. The variable S reflects this decision and takes a value of '1' if the shareholder sues and '0' otherwise. Filing a lawsuit comes at a cost 'a' to the shareholder. These litigation expenses are incurred regardless of the outcome of the suit. Finally, if the manager is discovered by law to have committed fraud, he pays a penalty 'p'.

[8] More generally, one can think of three levels (perfect, partial and no monitoring) or even a continuum of monitoring levels. In the continuous setting, different levels of monitoring give a manager differing opportunities to commit fraud.

The manager's payoffs and the shareholder's payoffs depend on their actions and are described in the table below. Payoffs in cells are the manager's payoff followed by the shareholder's payoff.

Shareholder's decision (S=1 indicates a lawsuit)

Manager's decision (F = 1 indicates fraud)		S =1	S =0
F =1		$\xi - p$, $-c - \xi + p - a$	ξ, $-c - \xi$
F =0		0, $-c - a$	0, $-c$

For any interesting conclusions, it must be assumed that $p-a \geq 0$, i.e. the shareholder gets a positive benefit by suing if the manager is discovered to have committed fraud. (If this is not the case, even having perfect knowledge of fraud will not induce the shareholder to sue as litigation is too costly.) The values of all variables including p, a and ξ are known to both parties. In this perfect information model, the shareholder can perfectly observe the action the manager has taken, i.e. whether F=0 or 1.

This model can be solved by backward induction. Consider first the case where M_L or weak monitoring is chosen.

i) The shareholder's decision at t=2, given she has chosen M_L earlier is as follows: Since F is perfectly observable, the shareholder's decision can be directly conditioned on it. The shareholder will sue (S=1) if fraud is observed (F=1) and not otherwise (S=0 otherwise).

If instead the shareholder had chosen M_H, the stronger monitoring structure, at time 0, then this would have precluded any fraud and thus the shareholder's choice at t=2 would always be not to sue (S=0).

ii) Assuming that $\xi - p \leq 0$ always (the manager is made to return all the money appropriated and perhaps even pay a penalty[9]), then the manager's decision at t=1 is not to commit fraud (F=0). This assumption implies that the manager never commits fraud in a perfect information world.

iii) Finally, the shareholder's decision at t=0 is always to choose the weaker monitoring structure, M_L, because outcomes are observable and litigation can always be selected in case fraud is observed to occur.

This benchmark perfect information model gives us the interesting result that in a world where fraud is observable, it is optimal not to impose any ex ante monitoring controls. Subsequent legal action conditioned on observing managerial misdeeds is the best course of action. If losing lawsuits is costly to managers, fraud is never an optimal choice. This benchmark tells us that we see governance and monitoring in the world only because shareholders can not completely observe fraud or cannot prove it beyond doubt in the courts; i.e. due to imperfect information conditions.

Imperfect information model

In a more realistic model, the value of F (the variable representing fraud) is not directly observed by the shareholders. Instead, what they observe is a noisy signal of

[9] This is quite consistent with our observations of real-world outcomes of litigation in which the fines often exceed the actual losses suffered by the shareholder. Another way of looking at this would be that instead of the manager directly compensating the shareholders for their monetary loss, the payment is made by the D&O insurer of the company to the shareholders, as is often the case in the real world. However the manager suffers reputational penalties at least equal to the monetary reimbursement. The effect of this separation is the same – the shareholder is compensated and the manager pays a penalty.

fraud, namely whether the firm has restated its financial statements or not. This is denoted by the restatement variable r which takes the value 1 or 0 based on F. If F =1, r is assumed to take the value of 1 with the probability θ and 0 with probability (1- θ). If F =0 then r takes the value 0 with probability θ and 1 with probability (1- θ). Note that the accuracy of the signal (θ) can take any value as long as it is informative.[10]

Actions of the manager and the shareholder are represented in extensive form in Figure 3.1. The final payoffs to manager and shareholder remain the same as before and are unaffected by the r variable. The decision boxes connected by a dotted line represent an information set. If a shareholder sees that a restatement has occurred (r=1), she cannot tell if the manager has committed fraud or not, and hence the shareholder's decision boxes following the two outcomes (F=1 and r=1) and (F=0 and r=1) are connected by a dotted line. The same applies to the shareholder's information set after no restatement has been observed.

As in the perfect information case, the shareholders' problem at t =2 is analyzed first, when M_L was chosen at t=0. (In the simple model, if M_H was chosen this eliminates all opportunities to the manager to commit fraud, and hence S =0 is always the optimal choice regardless of r.) I propose that there exist two mixed-strategy equilibria in this model which are described below.

Equilibrium 1: Sue if restate =1, mixed strategy if restate=0

[10] These conditional probabilities of a restatement do not need to be the same for F=0 and F=1. Without loss of generality, they are both assumed to be equal to in order to simplify the variable structure.

Denote the initial (unconditional) probability that a manager commits fraud by μ. Note that this is unconditional and unrelated to the value of r. Consider first the case where the shareholder sees that no restatement has occurred (r=0). In this case, suppose she chooses to sue (S=1) with a probability of α and not to sue (S=0) with probability (1-α). The posterior probability in terms of μ is Pr (F=1/r=0) = $\frac{\mu(1-\theta)}{\mu+\theta-2\mu\theta}$.[11] For this to be a mixed strategy equilibrium, the following two conditions must hold:

The shareholder is indifferent between choosing S=1 and S=0, given the mixed strategy employed by the manager. This gives us

$$\mu = \frac{a\theta}{p(1-\theta)-a(1-2\theta)} \qquad (1.1)$$

The manager is indifferent between choosing F=1 and F=0, given the mixed strategy employed by the shareholder. This gives us the value of α.

$$\alpha = \frac{\xi}{p} \qquad (1.2)$$

The shareholder's utilities from selecting S=1 and S=0 when r=1 are given below.

If S=1 is chosen, the shareholder's utility = $\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$ (-ξ+B–a) + (1-$\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$)(-a)

If S=0 is chosen, the shareholder's utility is zero. Given the value of μ in (1.1), the utilities under the two choices can be compared.

It is optimal to select S=1 when r=1. (1.3)

Equilibrium 2: Don't sue if restate=0, mixed strategy if restate=1

[11] See Appendix A for details of calculations.

Once again we consider first the mixed strategy case where the shareholder sees a restatement. In this case, suppose she chooses S=1 with a probability of β and S=0 with probability (1-β). The unconditional probability that the manager commits fraud is denoted by μ. We get Pr (F=1/r=1) = $\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$. For this to be a mixed strategy equilibrium, we need:

$$\mu = \frac{a(1-\theta)}{a(1-2\theta)+p\theta} \qquad (2.1)$$

$$\beta = \frac{\xi}{p} \qquad (2.2)$$

When the value of r=0 is realized, the result is a pure strategy response described below. The shareholder's payoffs are:

If S=1 is chosen, the shareholder's utility is $\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$(-ξ + p – a) + (1-$\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$) .(-a)

If S=0 is chosen, the shareholder's utility is zero. Given the value of μ in (2.1), the optimal choice on seeing no restatement is not to sue. (2.3)

Refinement of equilibria

In this imperfect information setting, both of the equilibria described above are subgame perfect. They can be further refined by evaluating whether they fulfill the necessary conditions to be sequential equilibria. Sequential equilibria have to fulfill the conditions of sequential rationality as well as consistency. By defining the shareholder's strategy as {S=0} off the equilibrium path, both equilibria described above will fulfill the conditions of sequential rationality and consistency and hence both are sequential

equilibria. Since payoffs are known to both parties and there are no hidden 'types' of players, equilibrium dominance cannot be used to eliminate either of them. However we can find the Pareto dominant equilibrium by comparing the probability of fraud μ in both equilibria from (1.1) and (2.1). Given our assumption that $p > a$, the instance of fraud is always higher in the first mixed strategy equilibrium where the shareholder plays a mixed strategy when $r = 0$. Thus equilibrium (2) Pareto dominates equilibrium (1). In other words, both these equilibria are equally likely and reasonable by most measures, but equilibrium (2) might be preferred in that it involves lower levels of both fraud and lawsuits.

Comparative statics

Given the expected payoffs of the shareholder as calculated above, we can now derive the initial decision of whether to choose high or low monitoring. Since firms are heterogeneous in monitoring costs, only shareholders of firms with a low monitoring cost c_i would choose to pay this upfront and ensure the absence of fraud. Shareholders of firms with a prohibitively high c_i would choose weak monitoring and sue later depending on the value of r they observe. In other words, if c_i were smaller than the expected cost of suing conditional on the signal r, the firm would choose to monitor rather than possibly make an erroneous judgment about the fraud (based on a noisy signal) and suffer a loss of litigation expenses.

In the empirical study that follows, I first investigate whether firms with one or more institutional blockholders and firms with a high proportion of independent directors are less likely to face litigation, consistent with the predictions of this model. Both institutional blockholders as well as independent directors are considered in the

literature to play a monitoring role in the company. I also identify differences in firm characteristics that could possibly account for firm-level variation in monitoring costs, and hence would determine the firm's choice between monitoring and litigation.

Chapter 4

Ex Ante Monitoring or Ex Post Litigation? Empirical Tests

Data description

The initial sample in this study consists of all the firms in the Compustat database over the 1997 to 2001 period. This sample is then further reduced to firms for which I could find board, analyst and compensation information as described below. Of this set of firms, I identify the firms that have experienced restatements and SCA suit firms from two separate lists.

The data on securities class action suits is from the Stanford Securities Class Action Clearinghouse and consists of all securities class action suits from 1996 to 2002, numbering about 1180 observations. The sample includes only securities class action observations that took place after the securities law reform (PSLRA[12]) in 1995 and before the implementation of the Sarbanes-Oxley Act of 2002. This restriction is in order to exclude any impact of regulatory and legal changes and to minimize structural breaks in the data.

[12] The Public Securities Litigation Reform Act (PSLRA 1995) was designed to discourage frivolous securities litigation. Among other measures, it transfers the defendant's legal fees to the plaintiff for claims lacking substantial legal and factual support and imposes limits on attorneys' fees.

The data on earnings restatements is primarily from the General Accounting Office database of restatements and consists of 963 firm-years (951 unique firms) that announced earnings restatements in the period from 1997 to mid-2002. According to the GAO report, the data "excludes routine restatements and those resulting from changes in accounting policies. It focuses only on restatements that indicate accounting irregularities, including so-called aggressive accounting practices, intentional and unintentional misuse of facts applied to financial statements, oversight or misinterpretation of accounting rules, and fraud."

On merging the restating and suit firms with the data available on Compustat and Execucomp, there remain 775 restating firms and 485 SCA suit firms in the sample. For test purposes, the characteristics of these firms are compared to those of the remaining firms in the database during the period 1997-2001.

The other variables used in this study are related to compensation, institutional ownership and board characteristics. The CEO compensation variables including proportion of shares held by management (*Shrownpc*), total current compensation including options (*TDC1*) and the Black-Scholes value of option grants (*BS_Valu*) are obtained from the Execucomp database. Institutional investors' holdings are obtained from 13f filings through the Thomson Financial database. For a smaller subset of the firms in this sample, the number of total and inside directors is available from the Compact Disclosure database.

The data are matched up as follows. Each restatement is matched with a lawsuit, if any, in the subsequent year. The use of only one year for matching is due to the federal limitations for class actions to be filed. The rule states that action must be brought within

one year after discovery of the facts constituting the violation (or within one year after they could reasonably have been ascertained). Given these rules, any subsequent or future restatement may not be closely related to the subject matter of the lawsuit. In order to compare corporate governance characteristics across restatement firms with and without securities class action suits in year t+1 (controlling for a restatement in year t), I examine compensation, board and monitoring data for the year prior to the year of restatement, or two years prior to the class action suit (year t-1 in this example). This is in line with my claim that governance and monitoring should be ex ante deterrents to earnings manipulation, and substitutes to the ex post measure of securities litigation. Using board composition and other variables in the year of the restatement can be flawed as the discovery of the earnings manipulation and subsequent restatement itself may cause substantial changes in board composition, executive compensation and other measures. At the same time, data from further back in the firm's history may have little or no impact on an event happening years later.[13]

Empirical tests

Each observation in the dataset is a combination firm-year, some of which have restatements announced, some of which have lawsuits announced, and many of which have neither restatements nor lawsuits. Of the 10,922 firm-years in this panel with Compustat asset data available, 771 (7%) have restatements and 485 (4%) have lawsuits. Thus, while restatements and lawsuits are not common, they are frequent enough to be a focus of study. Further, the restatements and lawsuits are not clustered in any one year,

[13] For instance, Strahan(1998) looks at measures of agency problems in firms in the year 1990 and then looks up to 8 years in the future for indications of shareholder-manager conflict in the form of lawsuits.

but are spread across the sample period. Figure 4.1 depicts the number of securities class action lawsuits by year. All data on securities class action lawsuits as well as settlements for this and subsequent tables are obtained from the Stanford Securities Class Action Clearinghouse, jointly maintained by Cornerstone Research. From the graph, there is a sharp peak in the number of lawsuits in the year 2001, due to a sharp increase in IPO-based litigation in that year. In order to avoid having results that are solely driven by this spike in lawsuits in 2001, I remove IPO-based lawsuits as far as possible from the sample before performing the tests described in this section.

The firms experiencing lawsuits in this study come from a variety of industries. However, and as might be expected, some industries are more prone to securities class action litigation. Table 4.1 examines this effect by listing the top ten industries which have experienced the maximum number of class action lawsuits in this sample period. It appears that high growth industries such as pharmaceuticals and technology are more prone to lawsuits. Securities class action lawsuits can differ in their exact allegation even though their overriding theme is very similar. Most lawsuits are triggered by improper disclosure by management leading to losses for shareholders. More details about the specific allegations in lawsuits are tabulated in Table 4.2, which lists the top ten most common allegations in lawsuits. From this table, it is clear that most lawsuits allege either errors of commission or omission with respect to disclosure by management: lawsuits might be triggered either by disclosing false information or by a failure to disclose relevant information, thereby resulting in losses to shareholders.

It is interesting to note that a very large proportion of these lawsuits never go to trial. They are instead settled out of court. Table 4.3 lists the ten largest settlements of

securities class action lawsuits in the sample period. The information is current as of February 13, 2006 and lists the largest settlements falling in the post-reform act period from 1996 to 2005. Though these high-profile cases are not the norm, these settlements represent a substantial portion of the company's total value, indicating that this is a very significant economic event in these companies' lives.

The fundamental research question I ask in this chapter is whether the existence of monitoring mechanisms such as blockholder monitoring or outside director monitoring is negatively related to the probability of being sued. As a first step towards answering that question, I present simple statistics regarding the correlation coefficient between lawsuits and some important variables in Table 4.4. Panel A lists correlation coefficients and P-Values of the correlation between litigation and different monitoring variables. Different measures of blockholder monitoring are considered, from a dummy variable to simply indicate the existence of a blockholder, to the number of block holders and the proportion of shares held by them. The total holdings by institutions and the number of institutional block holders are also considered. The director monitoring variables include the total number of directors, the number of independent directors and a dummy variable indicating whether the CEO is on the board of the company. The correlation coefficients indicate that all these different measures of monitoring are significantly negatively related to the subsequent occurrence of a lawsuit, thus providing a basis for further tests. The correlations in Panel A indicate that the incidence of securities lawsuits is negatively correlated with the existence of institutional blockholders in the firm. This result supports the hypothesis that securities litigation is a substitute for monitoring by blockholders of the firm. The negative relationship between the number of independent directors and

lawsuits also suggests that litigation is a substitute for monitoring by an independent board.

Panel B reports the correlations between firm characteristics and the occurrence of securities litigation. The CEO compensation variables are positively related to subsequent litigation. This supports the hypothesis that firms which pay their managers highly may lack sufficient internal controls and accountability and thus need lawsuits for ex post action. The correlations in Panel B also support the theory that complex firms (measured by *Intangible Assets* or *Stock volatility* respectively) are more likely to rely on litigation than on monitoring as a way to keep managers in check. The regression includes the natural *log of market capitalization* as a control for size, which is positively related to the likelihood of litigation. *Dividend yield* is expected to be negatively correlated with lawsuits in accordance with the model, as managers of dividend paying firms are not considered to be entrenched, and internal monitoring may be a cheaper option for such firms. However this hypothesis is not supported by the correlation coefficient between *dividend yield* and litigation which is statistically insignificant. Also, the proportion of *shares held by the CEO* and the proportion of *total CEO pay to sales* are not significantly correlated with the risk of litigation. This does not lend support to the theory that high CEO ownership works to align incentives and reduces the risk of shareholder litigation.

Any multivariate examination of the determinants of securities lawsuits must include a control for the occurrence of a financial restatement. Prior research (for example, Wu (2002), Palmrose, Richardson and Scholz (2004) among many others) has documented significant negative market reactions to restatement announcements. In order

to confirm that the restatements included in this sample are largely material restatements and have received negative responses from the market, I calculate cumulative abnormal returns (CAR) for the restating firms on the restatement announcement date. On average, restating firms experience a -6.5% buy and hold returns on the date of the restatement announcement and a -9.9% buy and hold return on the day following the announcement. CAR results are negative both in the days prior to and immediately after the announcement and all these results are significant at a high 0.1% level.

Table 4.5 identifies the characteristics of restating firms by reporting results from a logistic regression with the occurrence of a restatement as a binary dependent variable (taking the values 1 and 0) and firm characteristics as the independent variables. The regression includes controls for size and dummy variables representing sample years, as some years tend to have significantly more restatements. Although total compensation is not significantly different for restating firms as compared to other firms, *incentive-based pay* (computed here as the ratio of the Black-Scholes value of option grants to total compensation), both for the CEO and for the top 5 executives, is higher for the restating firms. This result is consistent with the findings of Beneish (1999), Kedia (2003) and Denis, Hanouna and Sarin (2005) who argue that stock options generate incentives to manipulate earnings. I do not find evidence of a significant linear relationship between *board size* (including log of board size), or the *proportion of independent directors* on the board, and the occurrence of a restatement. Other variables such as the size and independence of the board of directors and *dividend yield* are not significantly related to the probability of a restatement.

To continue to test whether lawsuits are an ex post substitute to effective ex ante monitoring, in this section I describe the results of further univariate testing. Wilcoxon mean-difference tests were performed that examine differences in average firm characteristics between firms with and without lawsuits.[14] Table 4.6 reports the results of these tests. A higher level of monitoring in firms that do not have lawsuits would support my hypothesis that lawsuits come into force in companies which have weak monitoring. Panel A provides tests of the substitutability of lawsuits and monitoring by listing the mean value of blockholder and director variables for firms in my sample which have and have not been sued. Column 1 contains the averages for firms that do not experience a lawsuit in the sample period. Column 2 consists of average values for firms that have experienced a lawsuit in the sample period. Firm-years in column 2 have a lower average incidence of monitoring by many measures, lending support to the hypothesis that litigation and blockholder monitoring are substitutes. These lawsuit firms have fewer analysts following them, fewer independent directors on their boards and a smaller proportion of their shares held by institutional blockholders. The last column gives the P-Value of the Wilcoxon test statistic, which is significant for most of these differences. Panel B of Table 4.6 compares mean values of firm characteristics between the same two sets of firms: firms which have not experienced lawsuits in this sample, and firms which have. Values of compensation variables are invariably higher in Column 2 compared to Column 1, supporting the hypothesis that firms with well-paid managers tend to choose litigation, possibly because it is harder for blockholders to monitor entrenched managers. *Dividend yield* is lower for lawsuit firms, again consistent with the hypothesis that

14 The Wilcoxon Z statistic is the normal approximation of the Wilcoxon Statistic and includes a continuity correction. Results without the normal approximation are identical.

managers of dividend-paying firms are easier to monitor internally. The lawsuit firms tend to be larger firms with predictably low return on assets in the year prior to the lawsuit. Finally, complexity of a firm's operations (proxied by the variables *PPE or tangible assets* and *Black-Scholes volatility* of the stock) appears to be positively related to the likelihood of a lawsuit, consistent with the prediction that complex firms are harder to monitor internally. However the standard deviation of analyst forecasts is lower for lawsuit firms as well, making lawsuits the preferred means of retaliation by shareholders when companies miss consensus forecasts, perhaps due to manipulation or malfeasance.

A further set of tests examine whether lawsuits and monitoring are substitutes, and if so, what factors determine the choice between them. Tables 4.7 reports the results of tests for a substitution effect between lawsuits and monitoring. A multivariate logit regression is used in which the dependent variable *Suit* is a binary-valued indicator that takes a value of 1 or 0. The regressions also control for the occurrence of a *restatement* and the results show that this factor is indeed an important determinant of the occurrence of subsequent lawsuits, as its coefficient is positive and significantly different from zero. The results also show that the *existence of a blockholder* is significantly negatively related to the occurrence of a suit. The number of *independent directors* on the board also appears to be significantly negatively related to the occurrence of a lawsuit. These results suggest that securities class action lawsuits are a substitute to monitoring by blockholders and independent directors. Lawsuits come into force as a disciplining mechanism in firms that do not have strong institutional blockholders who might be expected to monitor managers. Alternative specifications of the dependent variable for this test give very

similar results.[15] Finally, I test to see if the *number of analysts* that follow the activities of the company is seen as a form of monitoring and is negatively related to a subsequent lawsuit. However I do not find any evidence to support this claim. The regression results include year and industry dummy variables to control for these fixed effects. I also adjust the standard errors for clustering of variables due to multiple observations of the same firm through different years. The control variables include size and profitability measures.

The variables used in the logit regression analysis, like variables used in similar studies, are subject to endogeneity concerns. In a regression of litigation on blockholding measures, the error term may be positively correlated with the blockholding measure leading to biased estimates of OLS coefficients. Results of a Hausman test comparing OLS and Two Stage Least Squares (2SLS) estimates of this regression in Table 4.8 confirm that this is a valid concern. The Hausman test statistics indicate that 2SLS gives superior estimates of regression coefficients compared to OLS, both for the relationship between blockholder existence and litigation, and institutional holdings and litigation. To address this concern, I repeat the test conducted in Table 4.7 but using 2SLS procedures instead of logistic regression. I use an indicator variable that assumes a value of 1 if the company is a member of the S&P 500 index. This variable is used to instrument for monitoring by blockholders or institutional blockholders. An S&P 500 indicator is a good instrument for institutional holdings as institutions show a strong preference for S&P index stocks. Results of the first stage regression of institutional holdings on this S&P

[15] An alternative specification is to specify a multinomial choice regression by coding observations with lawsuits as '1', observations with an institutional blockholder as '3' and observations with both a lawsuit and an institutional blockholder as a '2'. Results remain unchanged.

membership variable are presented in Panel A of Table 4.9. The instrument is closely correlated with the existence of a blockholder, a high number of blockholders and a high proportion of institutional holdings. In the second stage, results of which are presented in Panel B, the occurrence of a lawsuit is regressed on this instrument and other controls. The significant negative coefficients -.1511 and -.0239 confirm the previous finding that litigation and institutional shareholdings are negatively related. Results from the 2SLS analysis confirm the findings from the logistic regressions, both in general and regarding the negative relationship between litigation and blockholder or institutional monitoring.

One of the predictions of the model concerns the probability of a lawsuit following a restatement. According to the model, a shareholder views a restatement very differently depending on whether it comes from a strongly or weakly monitored firm. A restatement from a strongly governed firm is attributed to error whereas a restatement from a weakly governed firm is more likely to be followed by a lawsuit. To test this prediction, I calculate the marginal effects of a restatement on the probability of a lawsuit. Table 4.10 reports the marginal effects of a restatement on the probability of a lawsuit, for firms with and without an institutional blockholder. This table provides a further test of the predictions of the theoretical model. Indeed, on average the incremental probability of facing a lawsuit when a firm restates its financial statements is lower for firms with at least one institutional blockholder (0.08) than for firms without institutional block holders (0.14). Similarly, the sample is also divided into firms that have an above-average proportion of institutional shareholders and those that have a below-average proportion of institutional shareholders. The proportion of institutional shareholders is calculated by dividing total shares owned by institutions by the total shares outstanding.

Once again, a restatement by a firm with higher than average proportion of institutional monitors appears to be viewed benignly (is less likely to be followed by a lawsuit) as compared to a restatement by a firm with less institutional shareholdings.

The results in Tables 4.9 and 4.10 provide a strong test of the primary predictions of the model developed in Section II. From Table 4.9, firms with institutional blockholders, or a higher proportion of shares held by institutions, are less likely to be sued. From Table 4.10, a restatement coming from a weakly monitored firm is more likely to be followed by a lawsuit than a restatement from a strongly monitored firm.

Additional evidence

In this section I report the results of robustness checks to test whether my hypothesis regarding the substitutability of monitoring and litigation is supported by any further evidence.

Do institutions monitor or vote with their feet? Institutional holdings pre and post lawsuits

The empirical literature examining the role of institutional investors has proposed two possible actions that institutions can take when they disapprove of the actions of the management of a company. The 'voting with their feet' hypothesis claims that institutions exit the stocks of companies whose practices they disapprove of. The monitoring hypothesis, in contrast, argues that institutions keep their shareholdings and try to improve the companies from within through activism. This prompts the question of whether institutions anticipate lawsuits (or the manipulation or fraud that was their cause) and sell their holdings in a stock, or if they stay and monitor to try and improve performance of the company.

In order to examine this, I run an event-time regression with the proportion of shares in a firm held by institutions as the dependent variable. This regression includes only firms which that experienced securities class action lawsuits, and its results are in Table 4.11. In this regression, I use dummy variables for event years where year 0 is defined as the fiscal year when the lawsuit was filed, year 1 as the fiscal year after the lawsuit, and so on. The regression coefficients on these year dummies provide information about whether institutions own more of the stock in these years or less. I also

control for calendar year effects using year dummies as some years have more lawsuits than others. Other controls include Fama-French industry dummies, and other factors that have been found to influence institutions' decision to buy a stock, including *Book to Market, Size, profitability, leverage, previous year stock returns*, and past *share turnover* (a proxy for liquidity). I used lagged *share turnover* two years prior to the year in question, so as to remove the effects of any immediate illiquidity caused by the lawsuit announcement.

I find that, after controlling for the common determinants of institutional shareholder buying behavior, institutions do tend to move out of a stock that is about to face litigation. The coefficients for year 0 and year -1 are significant and negative, indicating that institutions sell the stock in the year of the litigation announcement and in the previous year. However in subsequent years, there is no strong evidence of selling, indicating that institutions may try to stay and improve internal controls of firms.

Have lawsuits decreased after Sarbanes-Oxley?

If strong governance or monitoring is in fact a substitute to ex post litigation, then an external shock to governance requirements that led to improved governance should also lead to a lower level of litigation. The passage of the Sarbanes-Oxley Act of 2002 was just such an external shock, imposing stringent governance and disclosure requirements on all public firms. It is a natural question, then, whether there has been a substantial reduction in securities class action lawsuits since Sarbanes-Oxley was instituted.

Securities class actions as a whole, however, increased since Sarbanes-Oxley due to a variety of reasons. Independent regulatory changes prompted a big jump in IPO-related litigation in 2001 which has continued into the next few years. A large proportion of the litigation has also targeted private firms that were not affected under Sarbanes-Oxley. In order to identify and isolate the Sarbanes-Oxley effect, I use Stanford Clearinghouse data on allegations and company characteristics, and consider only public firms in this part of the study. Figure 4.2 shows the number of lawsuits against public firms in the years prior and subsequent to the Sarbanes-Oxley Act. The fitted polynomial line shows a drop in the number of lawsuits against public firms after 2002. I also test this more rigorously by controlling for the common determinants of lawsuits, and testing for a structural change in the number of lawsuits in 2002 using a Chow test. Results of this test are given just below the figure and show that a significant change occurred at around the time that Sarbanes-Oxley came into effect, and that the number of lawsuits has significantly dropped. This additional evidence supports the conclusions described in the previous section.

Monitoring and fraud

The third robustness check involves examining the relationship between monitoring and fraud. If monitoring by institutions and boards is a substitute for litigation, then it must have an effect on the likelihood of fraud in a company. If evidence exists to show that monitoring by institutions reduces the likelihood of fraud, and litigation is an after the fact attempt to compensate shareholders for their losses due to managerial fraud, then it is easy to see that monitoring and litigation are substitutes. Monitoring then would prevent fraud, whereas litigation is a response to fraud.

I examine this proposition by looking at a sample of 70 firms that have been a target of the SEC's Accounting and Auditing Enforcement Actions, or AAERs. These firms are matched with a control sample of 70 other firms which have not experienced either AAERs or lawsuits, and which are in the same 4-digit SIC industry code as the sample firms. The control firms are selected so that they are closest to the sample firms in size and profitability, among the firms in the same industry. Using this matched sample of firms, I run a conditional logit regression in which 'SEC fraud' is the dependent variable. Results of this regression are in Table 4.12. This is regressed on size, profitability, leverage, and the proportion of accruals that are discretionary (an indication of earnings manipulation), and the monitoring variables. I test whether monitoring by institutions, analysts or by independent boards prevent fraud. However, I do not find strong evidence of any of these variables predicting the occurrence of fraud. Part of the reason for this result may be the very small sample size, and hence the results on this front are inconclusive.

Taken as a whole, these tests provide convincing evidence that monitoring by institutional investors, analysts, and independent boards is a substitute to litigation. In the next section I analyze corporate disclosure, one of the mechanisms by which good monitoring may mitigate the agency problem and reduce the incidence of fraud and litigation.

Chapter 5: The Quality of Disclosure in Annual Reports and the Risk of Litigation

In this section I develop a measure to analyze the quality of disclosure in annual reports through content analysis of the words used in the document. Poor disclosure by this measure is then tested for its ability to predict litigation by shareholders. I also compare this measure with existing measures of poor disclosure described in the literature.

Inadequate or misleading disclosure is one of the most common allegations in securities class action lawsuits.[16] This allegation of poor disclosure is frequently coupled with other allegations such as misusing information through insider trading and is a major driver of securities lawsuits. Though this relationship between poor disclosure and litigation is well known to securities law practitioners, no attempt has been made to study or quantify it. An analysis of the types of disclosure associated with litigation would be helpful in setting standards for the information flow between managers and the shareholders they represent. Moreover such a measure of disclosure quality would also be useful in a general sense. The aim of good disclosure is to convey sufficient and timely information to shareholders, and securities litigation is most frequently a stated response to poor disclosure. Hence any measure of disclosure that can predict litigation is also a good general indicator of poor disclosure.

The language used in managerial communications to shareholders can be an important signal of managements' vision and outlook towards the future of the company,

[16] See Table 4.2 for the most common allegations in securities class action lawsuits. The top three allegations concern poor disclosure. Data is from the Stanford Class Action Clearinghouse database.

as well as their perspective on past events and current profitability. The psychology literature has developed methods to analyze words used in both speech and writing, through content analysis packages and techniques. These packages use built-in dictionaries to categorize words according to various measures such as '*tone*' or '*optimism*'. A document analyzed in this manner is given scores along different dimensions and an analysis of these scores can provide information about the document along various dimensions such as positive or negative tone, vague or factual words, and the proportion of numerical information and forward-looking information. In my study I use two specific content analysis software programs which are described below.

Content Analysis Software Programs

Linguistic Inquiry and Word Count (LIWC)

Linguistic Inquiry and Word Count (LIWC) is a text analysis software program designed by James W. Pennebaker, Roger J. Booth, and Martha E. Francis. LIWC analyzes text along different dimensions including the rate at which the authors of the document use positive or negative emotion words, self-references, big words, unique or repetitive words, strong and weak words, predictive and vague words, future and past-related words, and overstatement and understatement. LIWC contains over 70 built-in categories and allows words to be present in more than one category. Unlike some text and content analysis software, LIWC operates simply by counting the number of words in the document under each category. This methodology, while straightforward, also has

certain limitations. For example, the software is unable to distinguish between different meanings of a word, and cannot modify its results based on usage in a sentence.

LIWC has been used in a number of studies: to differentiate between lying and telling the truth, analyze dominance in a conversation, and even predict consumer buying patterns. However to my knowledge this is the first study to use LIWC to analyze company documents, specifically SEC filings.

General Inquirer

The General Inquirer (GI) is a text analysis program developed by Philip Stone at Harvard and made available for academic research purposes. It specializes in analyzing large quantities of text and outputs a matrix of "tag counts" for each of 180 categories, with separate rows of counts for each file processed.

A distinguishing feature of General Inquirer is its disambiguation routine. The software specializes in differentiating between alternate meanings of homonyms according to the context. For example, it distinguishes between "race" as a contest, "race" as moving rapidly, "race" as a group of people of common descent, and "race" in the idiom "rat race".

The General Inquirer also cautiously removes common regular suffixes so that one entry in a category can match several inflected word forms. A category entry can be an inflected word (for example, "swimming"), a root word ("swim" would match "swimming", if "swimming" is not a separate entry) or a word sense (for example, "swim#1") identified by the disambiguation routines of an inflected or root word form. Even though these disambiguation routines often require the Inquirer to make several

passes through a sentence, the Inquirer is designed to process large amounts of text in a reasonable amount of time.

Though the General Inquirer has been extensively used in the psychology and political science literatures, it is only recently beginning to be used for finance research. Tetlock (2005) used this content analysis technique to analyze news articles in the Wall Street Journal and their relationship to stock price and volume. The results of the study were that high media pessimism predicts downward pressure on market prices followed by a reversion to fundamentals, and unusually high or low pessimism predicts high market trading volume.

Data Description

The data used from this study consists of over 70,000 electronic 10-K filings of companies downloaded from the Securities and Exchange Commission's (SEC) EDGAR database. All publicly held companies are required to file annual financial reports or 10-Ks with the Securities and Exchange Commission (SEC). The complete annual report, in contrast, is sent to shareholders and typically contains the same information, presented in a better and more readable format, aimed at communicating information to the company. Since the basic contents of the 10-K and the annual reports are the same or have few differences, and since 10-Ks are easily downloadable and available on the SEC's website, that is the document selected for the purposes of this study.

The dates of these 10-K documents range from 1993, when EDGAR's online submission process became widely available, to 2005. 10-Ks of all companies that have ever filed electronically with the SEC are collected and used in the analysis. The dataset is free of survivorship bias because it includes all filings made by companies that existed

in any year of the sample. This means that even filings of firms that subsequently dropped out of the sample are present in the years that these firms existed and filed reports. Information such as filing date, filing type and CIK number (the identifying number given by the SEC to its filers) were extracted from the filing and filing headers. Separately, I extracted companies' fiscal year start and end dates from Compustat and matched this information with the filings in order to identify the appropriate fiscal year that each filing pertains to.

These 10-K filings were subsequently run through the text analysis programs described in the previous section; LIWC and GI. Both programs output the total wordcount of the document, the number of words per sentence, absolute wordcounts in each of the word categories, and category wordcounts scaled by the total words in the document. This dataset was then supplemented using profitability measures, leverage and other accounting information from the Compustat database and fiscal-year returns from Compustat PDE (Prices, Dividends and Earnings). Further, analyst forecasts and forecast error for each firm were added from the I/B/E/S database.

In order to predict lawsuits using content analysis measures of annual report disclosure, I combine this dataset with lawsuit, governance, institutional shareholding and blockholding data. The securities class action data is obtained from Securities Class Action Clearinghouse-Cornerstone Research database for the period 1996-2002 maintained jointly by Stanford Law School and Cornerstone Research. I obtain board and inside director information from IRRC and institutional blockholder information from Thomson Financial's 13f filings. In addition, I use the Gompers, Ishii and Metrick G-Index or Governance Index which is an index of the number of antitakeover provisions

in the company's charter. I also use the Entrenchment Index or E-Index defined by Bebchuk et al. (2004), which comprises a subset of antitakeover provisions from the Governance Index.

Though the annual report or 10-K document is only one of the means by which managers communicate with shareholders, it is generally accepted as the most important. For example, Botosan (1997) states "Although the annual report is only one means of corporate reporting, it should serve as a good proxy for the level of voluntary disclosure provided by a firm across all disclosure avenues. This is because annual report disclosure levels are positively correlated with the amount of disclosure provided via other media (Lang and Lundholm (1993)). The annual report is the focus of my disclosure index because the annual report is generally considered to be one of the most important sources of corporate information." As another example, Knutson (1992) states "At the top of every analyst's list (of financial reports used by analysts) is the annual report to shareholders. It is the major reporting document and every other financial report is in some respect subsidiary or supplementary to it."

Empirical Tests

Using content analysis categories to predict lawsuits

In this section I describe the empirical analysis that uses the quality of disclosure to predict securities class action litigation. Word categories from the content analysis packages GI and LIWC described in the previous section are used to develop measures of the quality of disclosure. Appendix B contains a basic description of the selected word categories along with some of their constituent words. The correlation coefficients among

the LIWC word categories are described in Table 5.1 and correlation coefficients among GI word categories are described in Table 5.2. In the first section of the analysis, I analyze whether firms experiencing lawsuits have fundamentally different disclosure patterns along word category dimensions. In order to do this, I examine average values of the proportion of words in different word categories, across two subsamples: a set of firms that have experienced lawsuits and firms that have not. Word categories that are relevant to business context were selected from the multitude of categories. Table 5.3 lists the results of this analysis. Panel A lists the differences in LIWC word categories and Panel B lists differences in GI word categories between the two subsamples. The last column in both panels contains the P-Value of the Wilcoxon statistic which tests whether the average values of word categories are significantly different between the two subsamples. These univariate results (from Panel A) indicate that 10-K reports filed by lawsuit firms (averaged over all the years in the sample, and not just the year in which they experienced a lawsuit) are on average longer documents, containing fewer numbers, more tentative words and less information about the past and present. Also from Panel B, we can see that lawsuit firms use fewer *weak* words (possibly indicating a tendency to make extremely confident claims that may not be borne out), and a tendency towards overstatement.

Results from a multivariate analysis of the word categories closely related to the likelihood of a lawsuit are consistent with the univariate results discussed above. I use the word categories from LIWC and GI, from an analysis of the annual report prior to the litigation year, to predict the likelihood of litigation controlling for some of the common determinants of lawsuits and for firm-specific accounting and stock returns variables.

Results of these logistic regressions are in Table 5.5. The dependent variable in these regressions is '*Suit*', an indicator variable that takes the value of '1' if a securities class action lawsuit has occurred in that particular year. The independent variables are content analysis word categories from LIWC and GI which are used to predict the occurrence of a lawsuit. The data is a panel data set of firm-years, and the standard errors of the regression coefficients have been adjusted to account for the repeated observations of the same firm that occur in the sample. All LIWC and GI wordcounts are scaled by the total document length, and in the case of GI wordcounts, by the document length of GI-recognizable words. It is also important to note that all the content analysis has been carried out on the 10-K report in the year preceding the litigation, if any. Since the data includes both the date of submission of the annual report to the SEC, and the date of the lawsuit (and the class period), the observations are carefully matched. Hence the content analysis may provide a leading indicator of the type of behavior that may have led to the lawsuit, if not a direct cause of the lawsuit. The table also lists parameter estimates and P-values of each estimate. Also listed are the marginal effects of a slight increase in the independent variable on the probability that a lawsuit will occur (the dependent variable). Each regression specification has controls for years and industries (using Fama-French industry definitions) in the sample, in order to remove any industry or time-specific effects. In addition, the regressions also include controls for *size* (the log of market value of the shares of the company) and *profitability* (current, lagged and future return on earnings of the company) which are often highly correlated with lawsuits. In addition to these controls, I also control for whether a *restatement* has occurred in the year prior to the lawsuit. This is necessary because lawsuits are often triggered by restatements of

earnings, in conjunction with poor disclosure. Finally, I also control for *stock returns* in the fiscal year prior to the lawsuit as share price drops are frequently triggers of litigation. The table lists marginal effects of a slight change in the regressor on the increase in probability of a lawsuit. In addition to the word categories, I also control for various governance and monitoring variables which are significantly negatively related to the occurrence of a lawsuit as discussed in the previous section.

I use various word categories from LIWC in order to predict a lawsuit. The total number of words used in the 10-K document (*LIWC_WC*) stands out as a strong predictor of lawsuits, since it is positively and significantly related to the likelihood of a subsequent lawsuit. Also, documents which contain substantial proportions of negative emotional words are more likely to be followed by litigation. The results using LIWC categories also indicate that companies whose documents which contain more numerical information are less likely to be sued. I also test whether the proportion of words referring to *past*, *present* and *future* events in a company are related to the likelihood of a lawsuit. Though the results relating to present and future words are inconclusive, the proportion of words in the document related to the past is significantly negatively related to the likelihood of a lawsuit. The results with numerical information as well as words related to the past indicate that more quantitative information, including information about past events in a company, is a hallmark of good disclosure, or disclosure that does not drive shareholders to litigation. In the next regression, I use GI word categories including the *document length*, *positive* and *negative* words, *strong* and *weak* words, and *overstatement* and *understatement* words. Again, the document length is a significant predictor of litigation risk. The results using the '*negative words*' category also

corroborate the earlier results with the LIWC '*negative emotional words*' category. Firms whose documents have a higher proportion of negative words are more likely to be sued in the following year. Documents with more '*weak*' words are less likely to trigger litigation, whereas documents with '*strong*' words are more likely do so. This could be because companies which make strong claims for the future are unable to follow through and may be susceptible to litigation. On the other hand, companies which make moderate or weak claims are less likely to face the consequences of not meeting predicted targets. Finally, I do not find significant results for predicting litigation risk using the '*overstatement*' and '*understatement*' word categories.

The coefficients on the control variables are consistent with their predicted effects. An accounting restatement in the previous year leads to an increased likelihood of a lawsuit, as does poor profitability in the prior year, measured by return on equity. Also predictably, poor stock returns in the previous year is positively associated with the likelihood of a lawsuit. At the same time, monitoring by blockholders and board members is not significantly negatively related to the probability of a lawsuit, after controlling for the effect of disclosure in annual reports. This could indicate that monitoring by block holders and independent directors prevents lawsuits by improving the quality of disclosure in annual reports.

The last regression uses all word categories from GI and LIWC, using only a single control for the total document length. Most of the results from the separate regressions are valid in the combined regression as well. Total document length and the proportion of numbers in the document are still extremely significant variables in the regression. The use of more *negative words* leads to a higher likelihood of lawsuits, even

controlling for the use of other types of words. Further, the use of weak words leads to a lower likelihood of lawsuits and the use of strong words is associated with a higher likelihood of a lawsuit.

Unconditional logit regressions on panel data with dummy variables to control for fixed effects may be prone to estimation problems in small samples with a large number of regressors. Though the sample used in this study is very large, I perform a robustness check using the conditional logit regression method proposed by Chamberlain (1980). Results of this conditional logit regression are reported in Table 5.6 and are largely consistent with the results of the unconditional logit discussed above. 10-K documents filed by lawsuit firms are generally longer and have less numerical information compared to documents filed by non-lawsuit firms. These documents also provide less information about past and future events in the company. One may generalize and say that these documents, though longer, are less informative compared to documents filed by non-lawsuit firms. I define a more rigorous measure of informativeness and describe the procedure of calculating it below.

Using principal components analysis to predict litigation

In this section I describe the empirical tests in which I use Principal Component Analysis of the word categories to predict litigation. Principal component analysis (PCA) is a mathematical procedure that transforms a number of possibly correlated variables into a (smaller) number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible.

There are three advantages of using Principal Components based on word categories instead of using the individual word categories themselves. First, principal components analysis reduces the dimensionality of the dataset. Since the two content analysis software programs together account for more than 250 word categories, the use of Principal Component Analysis reduces these to a more usable number of dimensions that can be used as regressors to predict litigation. Second, the use of Principal Components comprising all or a large number of the word categories eliminates the use of subjectivity in selecting relevant word categories out of the large numbers of categories available in both software programs. Third, this procedure enables the discovery of more meaningful underlying variables: in this case, a variable capturing the quality of disclosure. Principal Components Analysis is also more suited for the purposes of this study than other commonly used techniques such as principal factor analysis.[17]

Mathematically, principal components are determined by solving for the eigenvalues and eigenvectors of a square symmetric matrix with sums of squares and cross products. The eigenvector associated with the largest eigenvalue has the same direction as the first principal component. The eigenvector associated with the second largest eigenvalue determines the direction of the second principal component. The sum of the eigenvalues equals the trace of the square matrix and the maximum number of eigenvectors equals the number of rows (or columns) of this matrix.

17 Where the purpose of principal component analysis is to derive a small number of linear combinations (principal components) of a set of variables that retain as much of the information in the original variables as possible, the purpose of common factor analysis is to explain the correlations or covariances among a set of variables in terms of a limited number of unobservable, latent variables. The latent variables are not generally computable as linear combinations of the original variables. In common factor analysis, it is assumed that the variables are linearly related if not for uncorrelated random error or unique variation in each variable; both the linear relations and the amount of unique variation can be estimated.

The principal components analysis of the word categories in this dataset yielded more than 25 components with nonzero eigenvalues. Figure 5.1 is a scree plot of the eigenvalues of each component, which provides information about the proportion of the total variance explained by each principal component. Based on the scree plot, I select the first 11 principal components to use in the analysis. The variance of the 11 selected principal components explains a total of 70 percent of the variation in the original word categories themselves.

More descriptive statistics about these principal components and the variables that comprise them are listed in Tables 5.7 and 5.8. Based on the word categories that comprise them, each component has been given a descriptive name. Thus the principal components are titled *wordcount, punctuation, wordiness, informativeness, intangibles, factual references, predictive words, positive words and small words and numbers*. Table 5.8 lists descriptive statistics including the means and standard deviations of these components.

In the next step, the occurrence of a lawsuit is regressed on the components described above, after controlling for some of the common determinants of lawsuits. The procedure used is Chamberlain's conditional logit model, which gives efficient estimates of the regression coefficients for panel data. The results of this principal component analysis are in Table 5.9. These results show that the fourth principal component, entitled '*informativeness*', is a significant robust predictor of lawsuits; since firms that provide informative disclosures in their 10-K documents are less likely to be the targets of litigation. The results hold even after controlling for current and lagged profitability, stock returns, governance and monitoring characteristics, and other principal components.

Comparing Principal Components with S&P Transparency and Disclosure Scores

In Chapter 2, I discuss different measures of transparency and disclosure that have been proposed in the literature. Some examples of these measures are the S&P Transparency and Disclosure score, the AIMR score, PIN or the Probability of Informed Trading and Botosan's (1997) and Francis, Nanda and Olsson's (2005) disclosure scores. Of these, I compare my disclosure *informativeness* measure with the S&P transparency and disclosure score because it is one of the most comprehensive scores awarded based on a systematic analysis of more than 98 different attributes. In addition, many of the other measures have limited applicability, are not available in my sample period (for example, the AIMR score was discontinued in 1995), or were awarded to a very small sample of companies.

I test the effectiveness of Principal Components in predicting litigation by comparing it to another well known measure of disclosure and transparency, namely the Standard and Poor's Disclosure and Transparency score.

First described for US companies in the paper by Patel and Dallas (2002), the S&P Transparency and Disclosure (T&D) rankings study was then broadened to encompass companies in several countries in Asia and Europe, which were ranked according to the method used for US companies. The aim of the S&P T&D rankings was to answer questions such as the following: Which companies provide the most extensive disclosure in their basic corporate filings? Which companies disclose above and beyond what the law requires? Are there significant differences among the T&D practices of various regions and within regions? What is the significance of the T&D rankings? And

how can T&D be improved? In order to assess T&D practices, the study identifies 98 disclosure items, classified into three broad categories: (1) Ownership structure and investor rights, (2) Financial transparency and information disclosure, and (3) Board and management structure and process.

In the US, Standard and Poor's awards two different ranks: the Composite rank and the Annual Report rank. The Transparency and Disclosure (T&D) rankings on a composite basis include, in addition to the annual reports, the regulatory-driven 10-K and proxy statements. This S&P disclosure score is based on 35 accounting items, 35 governance structures and 28 ownership structures. While U.S. composite disclosure levels were consistently high, disclosure levels based on annual reports alone were found to be much more variable.

While comparing rankings across countries, S&P found that the U.S. (composite basis) and the U.K. demonstrate the highest levels of disclosure globally, with Latin American and Asian emerging markets comparing the least favorably. Even where detailed disclosure is mandated by law, companies exercise much discretion about what they disclose, with some companies demonstrating higher discretionary disclosure standards compared with local norms.

The significance of the U.S. T&D rankings was also explored by Patel and Dallas (2002) by comparing Standard & Poor's rankings with factors affecting the cost of capital. The authors found that companies with higher T&D rankings (on both an annual report and composite basis) have lower market risk. In addition, companies with higher T&D rankings based on annual reports alone tend to have higher price-to-book ratios. They conclude that companies can lower the cost of equity capital by providing higher

transparency and disclosure. These findings were independently confirmed by Cheng, Collins and Huang (2003) who conclude that the S&P T&D rankings did provide new information to the markets along various dimensions, and that firms with large differences in disclosure levels across SEC filings were received unfavorably by investors.

I compare my principal components analysis of informativeness with the S&P T&D rankings for several reasons. The aim of the principal component analysis of the content analysis word categories was to come up with one or more measures that capture the disclosure quality of annual reports. The aim of the S&P rankings are very similar; the annual report ranking is a measure of the disclosure quality of the information contained in the annual reports whereas the composite rank is a measure of the degree of transparency of the company as whole, based on different types of public documents issued. However the S&P rankings are derived after considerable analysis of the written text and financial information disclosed by the company in its documents. These rankings are arrived at through a proprietary evaluation that incorporates the criteria that Standard & Poor's Governance Services uses in its interactive corporate governance scoring service and hence requires considerable human input. In contrast, the content analysis of word categories, and the subsequent principal component analysis (including the creation of the '*informativeness*' component) is meant to be easily replicable using publicly available data and software programs. In fact, the entire analysis can be automated. Both the GI and LIWC software programs are available free, and given the constituent word categories of each principal component (in terms of GI and LIWC word categories), it is easy to calculate the value of the '*informativeness*' component and other components for

a given document. The annual reports on which the results are based are also publicly available on the SEC's website.

Hence, if my analysis is found to have better predictive power for litigation over the S&P component, then firms exhibiting poor disclosure which is likely to lead to litigation can be identified by investors using the process described, with the publicly available information at the time the annual report is released. Table 5.10 lists correlations between my principal components, the two S&P Transparency and Disclosure ranks (annual report rank or AR Rank and the Composite rank), and litigation. P-Values listed below each correlation coefficient measure its significance. The *informativeness* component is not significantly related to either of the S&P rank variables. From a univariate comparison of the correlation coefficients of the principal components and S&P ranks with the litigation variable, it is clear that the *informativeness* principal component is strongly and significantly negatively related to lawsuits with a correlation coefficient of -0.1087. The S&P ranks are not significantly correlated with the occurrence of a lawsuit. Table 5.11 takes this analysis further in regressing the occurrence of a lawsuit on both the *informativeness* principal component as well as the S&P ranks. The regression includes controls for various common determinants of lawsuits including size, profitability, past stock returns and the occurrence of a restatement. As expected, a restatement positively predicts a subsequent lawsuit, as does poor stock returns. Neither the composite S&P composite rank nor the AR rank is a significant predictor of lawsuits. However the *informativeness* measure, principal component 4, is strongly negatively related to litigation. Poor informativeness is associated with the risk of shareholder litigation. The results hold even after controlling

for monitoring and governance measures, as well as the other principal components. The results show that the *informativeness* component is a much better predictor of lawsuits than either of the two S&P transparency and disclosure ranks.

To conclude, in this chapter I discussed the results of several layers of tests of the relationship between disclosure quality and litigation. First, I examined the relationship between content analysis word categories and litigation, and found that several dimensions of the content of 10-K reports are related to the risk of subsequent litigation. For example, a document that has more numerical information, more information related to the past and present, and fewer strong words or bold claims is less likely to be followed by litigation.

In order to combine these different aspects of the information contained in 10-K reports and combine them into one single measure of disclosure quality, I perform principal components analysis of these word categories. This results in an '*informativeness*' principal component that captures these aspects (numerical information, information related to past and future events, and factual content) and is found to be significantly negatively related to the risk of a subsequent lawsuit.

I test this measure of disclosure quality by comparing it to S&P Transparency and Disclosure ranks (T&D ranks), which are constructed from an in-depth analysis of annual reports and company statements. My *informativeness* measure is found to be a superior predictor of litigation. As the primary purpose of managerial disclosure is to provide information to satisfy shareholders, disclosure that predicts litigation can be considered an objective benchmark for poor disclosure. In the next section I discuss determinants of informative disclosure as measured by this *informativeness* variable.

Chapter 6: Does Good Governance Influence Disclosure?

In this chapter I discuss my theory and tests of the hypothesis that good governance and monitoring leads to better and more informative disclosure. I examine different aspects of governance including board size and composition, executive pay practices, antitakeover defenses, and monitoring by institutional shareholders and analysts in order to determine whether these factors can affect the quality of disclosure by a company. Specifically, I test whether these measures of governance systematically affect the types of words used in the annual reports to enhance understanding and clarity. The aim of these tests is to examine whether the informativeness of an annual report (measured using the *informativeness* principal component discussed in the previous chapter) is an outcome of the type of governance in place in the company and the degree of monitoring by institutional investors, analysts and board members.

The basic test in this section involves regression of the '*informativeness*' measure on governance and monitoring variables. In order to carry out these tests, I divide the governance and monitoring variables into two categories: the predetermined and the jointly determined. Governance variables such as board size and composition, the number of antitakeover defenses, and the executive pay practices are the 'predetermined' variables, or the ones that may be considered as exogenous in the regression equations. Antitakeover defenses are built into the charter of the company and are by no means affected by the type of disclosure in the company. Similarly, board size and composition, and the size and composition of CEO pay, are externally determined prior to the instance of the annual report-based informativeness measure. These variables are also slow-

moving and do not change much from year to year. It is hard to imagine how the quality of disclosure in the annual reports may affect these variables; whereas there is clear causality in the opposite direction. I use OLS regressions with panel data for each company and year, to regress the dummy variable indicating lawsuits on this spectrum of governance variables, after controlling for the common determinants of lawsuits. These regressions reveal that governance variables do in fact affect the quality of disclosure and strong governance is linked to better disclosure practices as well.

Tests involving the second category of variables, the 'jointly determined' monitoring measures, are structured differently. Institutional holdings and analyst coverage fall under this category of variables for which there is considerable doubt about the direction of causality. Good disclosure may attract institutional investors to hold a stock; alternatively, institutional shareholders may elicit good disclosure by the management. Similarly, a large analyst following may reduce the information asymmetry sufficiently so that companies do not need to follow extensive disclosure practices. Alternatively, a large analyst following may pressure management into disclosing more information. Given this ambiguous causality, an OLS regression does not provide the best test of these effects. Instead, I perform a two-stage least squares regression using instrumental variables for analyst coverage and institutional holdings. Results from the two stage least squares regression show that companies with high institutional shareholders and analyst coverage do have more informative disclosure practices.

Data Description

The dependent variable in these regressions is '*Informativeness*', a principal component extracted from the content analysis of word categories obtained from the

LIWC and GI packages. The word categories comprising this measure are described in Table 5.7. More details about the extraction of this component and the procedure of principal component analysis are available in Chapter 5.

Several governance and monitoring variables are used as controls in these regressions. *Board size* and the *number of independent directors* on the board are obtained from the IRRC director database. *G-Index*, a measure of shareholder rights, is also obtained from the IRRC database. *G-Index* is a number from 0 to 24 and its value indicates the number of antitakeover provisions in the firm's charter from a list of possible provisions. A high value of *G-Index* denotes more antitakeover provisions and hence weak shareholder rights to take over control of the company. Bebchuk et al. (2005) constructed their own version of an index of antitakeover provisions called *E-Index* or Entrenchment Index. This measure is based on 6 of these 24 antitakeover provisions that they consider are most relevant: four constitutional provisions that prevent a majority of shareholders from having their way (staggered boards, limits to shareholder bylaw amendments, supermajority requirements for mergers, and supermajority requirements for charter amendments), and two takeover readiness provisions that boards put in place to be ready for a hostile takeover (poison pills and golden parachutes). A high value of the E-Index indicates a more entrenched manager. I use both the *G-Index* and the *E-Index* in this study. CEO compensation information is from the Execucomp database. Institutional investor holdings are obtained from the 13f filings in the Thomson institutional investor database. Analyst coverage is from I/B/E/S. I also use several control for the common determinants of lawsuits such as market capitalization (a proxy for size), past and current profitability, leverage, and other accounting information from

Compustat, fiscal year stock returns from CRSP, and restatement data from the GAO database. All regressions include controls for year and industry. I use the Fama-French industry codes to control for industry effects.

Regressions using predetermined governance variables

Table 6.1 presents results from regressions of the informativeness variable on the first category of predetermined governance variables, and controls. Controls include size (the natural log of the total market capitalization of the company), profitability (log of the return on equity), the occurrence of a restatement, past fiscal year stock returns, year dummies and Fama-French industry dummies. As is to be expected, I find that poor past stock returns and poor past profitability appear to be followed by less informative 10-K filings. In addition to this, the positive significant coefficient on the proportion of independent directors indicates that a higher proportion of independent directors on the board of directors leads to more informative disclosure practices. A larger board of directors also leads to significantly more informative disclosure practices. In a separate regression, I also control for the first principal component: *document length*. A longer document is related to poorer disclosure (as measured by *informativeness*). The results with board size and composition persist despite adding new controls. I also include the *G-Index* in the regression and find that it has a significant negative impact on the quality of disclosure. This is consistent with more antitakeover measures resulting in entrenched CEOs who do not disclose information adequately. These results hold even after controlling for *discretionary accruals* which is an accounting measure of disclosure quality. Higher *discretionary accruals* are perceived by the accounting literature as a

strong indication of earnings management. It appears that firms that engage in earnings management are also likely to have poor disclosure practices.

I also examine the effect of CEO compensation variables on the quality of disclosure. The compensation variables examined here are the total pay of the CEO and the proportion of the total pay that is incentive-based. *Incentive pay* is the sum of the Black-Scholes value of options granted to the CEO and the value of any restricted stock grants, divided by the total pay. Using different specifications, I do not find any indication that the total amount of the CEO's compensation, or the proportion of the compensation that is incentive-based, affects the informativeness of disclosures by the company in the 10-K statements. The results in Table 6.1 are significant because disclosure quality is strongly influenced by the proportion of independent directors on the board and the degree of entrenchment of the top management measured by G-Index. This result indicates that CEOs who are governed by an independent board tend to disclose information in a timely and accurate manner in the annual reports to the satisfaction of shareholders. Since the *informativeness* measure is strongly related to the risk of subsequent litigation, this result provides a way that companies may reduce their exposure to shareholder lawsuits. Setting up an independent board might provide oversight over their management and elicit disclosure practices that are desired by shareholders.

Regressions using jointly determined monitoring variables

Table 6.2 reports the results of a two stage least squares regression in which I study the effects of institutional holdings and analyst coverage on disclosure informativeness using instrumental variables. The two stage least squares procedure is

used instead of a regular OLS regression in order to control for the endogeneity between disclosure quality and both institutional holdings as well as analyst following.

In Panel A, I report the results of a first-stage regression on which different measures of institutional ownership are regressed on an S&P membership variable, which is used as an instrument to predict institutional ownership and thus institutional monitoring. Then the fitted values from this first stage regression are used as a regressor in the second stage regressions in which the informativeness of a company's disclosure is regressed on this variable and other common determinants of lawsuits. In the second stage regressions reported in Panel B, I examine the determinants of informativeness by using it as the dependent variable in the regression. Different proxies are used for monitoring by shareholders including the existence of a blockholder, the number of blockholders and the proportion of shares held by institutions. I also include monitoring by analysts, measured by the number of analysts who cover the shares of a company. The coefficients on these different proxies for institutional blockholder monitoring and monitoring by analysts are all positive and significant, indicating that monitoring by both institutions and analysts is strongly related to the likelihood of good disclosure. The results hold after controlling for size, profitability and other measures that may influence disclosure.

The results from these two tables point to some corporate actions that may promote good disclosure practices: independent boards, monitoring by analysts and monitoring by institutional blockholders. It is interesting to note that these characteristics, widely accepted as good corporate governance practices, are related to good corporate disclosure as well, and adopting them can reduce the risk of litigation for a company.

However, other practices such as performance-based compensation do not appear to have an effect on disclosure quality.

In conclusion, this chapter describes the tests conducted to examine the determinants of the informativeness of disclosure quality in annual reports. 10-K documents of a company filed with the SEC are run through the GI and LIWC content analysis software programs, and their scores along various content dimensions are converted to a single loading on the principal component '*Informativeness*'. Using Two-Stage least squares estimation to control for possible endogeneity, I find that companies with independent boards, a large analyst following and shares held by institutions and blockholders are more likely to have good disclosure practices. The results in these regressions imply shareholders can induce better disclosure from management through oversight by boards, analysts and institutions, and not through managerial compensation. This result reveals yet another benefit of monitoring by boards and institutional blockholders that the academic literature has not examined closely. Monitoring not only has the direct effect of shareholder oversight of managerial actions, but it also has the indirect effect of inducing good communication flow between managers (agents) and shareholders (their principals).

Chapter 7: Conclusion

This dissertation examines the substitution effects between monitoring and lawsuits; and the mechanism by which strong governance and monitoring mitigate the agency problem: disclosure. The firm is characterized by an asymmetry of information between managers and shareholders, one manifestation of the agency problem. Insufficient, untimely or uninformative disclosures by management exacerbate this agency problem. This may result in dissatisfied shareholders who resort to litigation in order to reclaim the losses they have suffered due to misleading or false statements by management. Improper disclosure by management may also be a conscious attempt by management to cover their tracks from fraudulent or inappropriate activities. Hence good governance and monitoring may enhance the quality of disclosure which prevents shareholder litigation.

I examine this hypothesis in several parts. The first section examines the substitutability of litigation and monitoring (by institutional shareholders and independent boards) and finds that monitoring is a substitute for litigation. The theoretical model in Chapter 3 describes the shareholders' decision as one of strong internal monitoring or corrective measures such as litigation. Shareholders of a company have the option to set strong internal controls such as independent boards, good compensation practices, and analyst coverage. These measures help keep managers accountable to shareholders and keep their actions transparent, reducing the likelihood of securities litigation. The tests in Chapter 4 verify the empirical implications of this theoretical model. Using both regular logistic regressions as well as two stage least

squares regression, I confirm that monitoring by independent boards and institutional shareholders is a substitute to securities litigation. The two stage least squares regression controls for the endogenous nature of monitoring by institutional shareholders by using membership in the S&P 500 as an instrument for it. This instrument is also found to be significantly negatively related to the occurrence of a lawsuit even after controlling for the common determinants of lawsuits.

The next essay in Chapter 5 examines poor disclosure as a trigger to litigation, and develops a measure to quantify the informativeness of disclosure in 10-K filings. In this section, I analyze text in 10-K filings with the SEC using content analysis packages, and use the resulting word counts in different word categories to predict lawsuits. My findings are that filings which are long but contain a smaller proportion of informative words such as numbers, past and future related words are more likely to precede shareholder lawsuits. I use both unconditional logit regressions and Chamberlain conditional regression models to verify my results. In order to consolidate the different word categories and develop a single measure that captures the quality of information, I use principal components analysis of the word categories given by the LIWC and GI content analysis packages, and capture the different dimensions of the word categories in 11 orthogonal principal components. The principal component '*informativeness*' is found to be a strong predictor of the risk of securities litigation in the subsequent year. In logit regressions predicting a lawsuit using various principal components, *informativeness* has a significant negative coefficient, even after controlling for other principal components, accounting measures and governance and monitoring mechanisms in place in the company. Here again, I perform a robustness check using conditional and unconditional

logit regression methodologies, which give consistent results. Further, I compare this measure of informativeness with another measure of transparency and good disclosure that has been proposed in the literature, the S&P transparency and disclosure rank. These ranks are awarded based on comprehensive corporate governance analysis and rating, and the annual report rank is based on disclosure of corporate governance practices, board structure and other corporate information in the annual reports. These ranks are not closely related with the content analysis word count results from the 10-Ks, or with the principal components. In a horse race between the principal components and the S&P ranks, the *informativeness* principal component is better at predicting litigation, compared to the S&P rankings. This predictive power holds even after controlling for other common determinants of lawsuits, and for governance and monitoring mechanisms. Surprisingly, a high S&P rank for transparency does not work to prevent litigation. It is unrelated to litigation risk after controlling for factors that predict litigation. On the other hand, the *informativeness* principal component is a good negative predictor of the risk of litigation because companies with more informative annual reports are much less likely to be sued by shareholders.

The last essay in Chapter 6 examines the determinants of informative disclosure in 10-K filings by companies. The hypothesis tested is whether strong governance and monitoring lead to more informative disclosure. I separate the governance and monitoring mechanisms into completely exogenous mechanisms and possibly endogenous mechanisms and estimate the effect of each group separately. Using logistic regressions, I conclude that exogenous variables like the number of antitakeover measures and the proportion of non-independent directors on the board are significantly negatively related

to the likelihood of a lawsuit. In order to establish the impact of the possibly endogenous variables like blockholder monitoring and analyst monitoring, I use a two stage least squares approach using an instrumental variable, with similar results. Strong monitoring by institutional shareholders and analysts results in a better quality of disclosure in annual reports.

This research has important implications for research in governance and corporate disclosure. Content analysis programs that are now beginning to be used in finance are a great tool to examine the nature of managerial disclosures, and offer the added advantage of being programmable and automated. Thus the disclosure quality index I create can be easily computed from publicly available filings and provides an effective description of different facets of the annual report. The use of this disclosure quality index to predict possible fraud and litigation can represent a significant addition to the body of work that examines disclosure, in the finance, accounting and legal fields. The analysis of the determinants of informative disclosure can help investors create conditions in firms that are conducive to these types of disclosure by management.

To summarize, I find that monitoring by independent boards, analysts and institutional shareholders helps to enhance the informativeness of corporate disclosure which can result in less litigation. Since litigation is a costly mechanism for shareholders, this study points to alternatives that may be equally effective at a lower cost. The results of these studies underscore the importance of monitoring and activism by shareholders as well as of good governance structures within a company, in order to mitigate the agency problem and avoid costly shareholder litigation.

Appendix A

Imperfect information model

Equilibrium 1: Sue if restate =1, mixed strategy if restate=0

Probability that a manager commits fraud = μ. When the shareholder sees that restate=0, she chooses s=1 with a probability of α and s=0 with probability (1-α).

We can now get the following posterior probability in terms of μ as:

$$\Pr(F=1/r=0) = \frac{\Pr(r=0/f=1).\Pr(f=1)}{\Pr(r=0/f=1).\Pr(f=1)+\Pr(r=0/f=0).\Pr(f=0)}$$

$$= \frac{\mu(1-\theta)}{\mu+\theta-2\mu\theta}$$

The probabilities with which the agents undertake each action can be incorporated into the payoff table as follows:

		Shareholder's decision		
	r = 0 case	S=1 with prob. α	S=0 with prob. (1-α)	
Manager's decision	F=1 with prob. $\frac{\mu(1-\theta)}{\mu+\theta-2\mu\theta}$	ξ – p, -ξ + p – a	-ξ, - ξ	
	F=0 with prob. $\frac{\theta(1-\mu)}{\mu+\theta-2\mu\theta}$	0, - a	0, 0	

Given these probabilities, there are 2 conditions both of which have to be fulfilled for this to be a mixed strategy equilibrium.

The shareholder is indifferent between choosing S=1 and S=0, given the mixed strategy employed by the manager.

$$\frac{\mu(1-\theta)}{\mu+\theta-2\mu\theta}\ (-\xi + p - a) + (1-\frac{\mu(1-\theta)}{\mu+\theta-2\mu\theta})(-a) = \frac{\mu(1-\theta)}{\mu+\theta-2\mu\theta}\ (-\xi) + (1-\frac{\mu(1-\theta)}{\mu+\theta-2\mu\theta})(0)$$

This gives us the condition

$$\mu = \frac{a\theta}{p(1-\theta)-a(1-2\theta)} \qquad (1.1)$$

The manager is indifferent between choosing F=1 and F=0, given the mixed strategy employed by the shareholder.

$$\alpha(\xi - p) + (1-\alpha)(\xi) = \alpha(0) + (1-\alpha)(0)$$

This gives us the value of α.

$$\alpha = \frac{\xi}{p} \qquad (1.2)$$

To completely define the equilibrium, we also need to specify what happens when the value of r=1 is realized. The value of μ is the same in this case, as 'μ' denotes the unconditional probability of committing fraud. We can thus derive the posterior probabilities once again in this case as follows:

$$\Pr(F=1/r=1) = \frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$$

If the shareholder chooses S=1, her total utility is

$$\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}\ (-\xi + p - a) + (1 - \frac{\theta\mu}{1-\theta-\mu+2\theta\mu})\ (-a)$$

If she chooses S=0, then her utility is zero.

Substituting the value of μ from (1.1), we get that the utility to the shareholder from choosing S=1 is greater than the utility to the shareholder from choosing S=0. Thus under this equilibrium, if the shareholder sees r = 1, in equilibrium she will choose S=1.

(1.3)

Equilibrium 2: Don't sue if restate=0, mixed strategy if restate=1

Once again we consider first the mixed strategy case where the shareholder sees that restate=1. In this case, suppose she chooses S=1 with a probability of β and S=0 with probability (1-β). Denote the unconditional probability that the manager commits fraud by μ. We get

$$\Pr(F=1/r=1) = \frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$$

The table of payoffs can be rewritten as:

		Shareholder's decision	
	r = 1 case	S=1 w/ prob β	S=0 w/ prob. (1-β)
Manager's decision	F=1 with prob. $\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$	$\xi - p$, $-\xi + p - a$	ξ, $-\xi$
	F=0 with prob. $1 - \frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$	0, -a	0, 0

The 2 conditions both of which have to be fulfilled for this to be a mixed strategy equilibrium are given below.

The shareholder is indifferent between choosing S=1 and S=0, given the mixed strategy employed by the manager.

$$\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}\ (-\xi+p-a) + (1-\frac{\theta\mu}{1-\theta-\mu+2\theta\mu})(-a) = \frac{\theta\mu}{1-\theta-\mu+2\theta\mu}(-\xi) + (1-\frac{\theta\mu}{1-\theta-\mu+2\theta\mu})(0)$$

This gives us the condition

$$\mu = \frac{a(1-\theta)}{a(1-2\theta)+p\theta} \tag{2.1}$$

The manager is indifferent between choosing F=1 and F=0, given the mixed strategy employed by the shareholder.

$\beta\ (\xi - p) + (1-\beta)(\xi) = \beta\ (0) + (1-\beta)\ (0)$

This gives us the value of β as $\beta = \frac{\xi}{p}$ (2.2)

When the value of r=0 is realized, we again get a pure strategy response described below. The shareholder's payoffs are:

If S=1, utility to the shareholder is = $\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$ (-ξ + B – a) + (1 - $\frac{\theta\mu}{1-\theta-\mu+2\theta\mu}$)(-a)

If S=0, utility to the shareholder is zero.

Substituting the value of μ from (2.1), we get that the utility to the shareholder from choosing S=1 on seeing a signal r=0 is less than zero. Hence the shareholder will choose S=0 if she sees r = 0. (2.3)

(2.1), (2.2), (2.3) together define this equilibrium.

Appendix B

Sample constituents of word categories

Some content analysis word categories used in the analysis are listed here along with selected words that constitute these categories. The number of words constituting each of these categories is also listed in parenthesis. These word categories are from the content analysis software General Inquirer. The General Inquirer builds word categories from four sources: (1) the Harvard IV-4 dictionary, (2) the Lasswell value dictionary, (3) several categories recently constructed, and (4) "marker" categories primarily developed as a resource for disambiguation, but also available to users. There are two word categories each for the 'Positive' and 'Negative' words; these do have several overlapping words but differ in the dictionaries they use. Certain words may have several meanings depending on the context, these are assigned 'word senses', and different word senses may fall under different word categories. For example, the words 'aid', 'allow' and 'ally' on this list may be used in different senses: noun or verb. Their usage in a sentence affects what word category they are assigned to.

Positive (1914 words)	**Negative** (2293 words)	**Positive 2** (1046 words)	**Negative 2** (1165 words)	**Strong** (1902 words)	**Weak** (756 words)
Abide	Abandon	Advantageous	Adverse	Adept	Affliction
Able	Abate	Affection	Afflict	Adeptness	Anxiety
Abound	Abdicate	Affectionate	Affliction	Adjust#1	Apathetic
Absolve	Abhor	Affirm	Against	Administer	Apathy
Absorbent	Abject	Agree#1	Aggravate	Administration#1	Apologetic
Absorption	Abnormal	Agreeable	Aggression	Administrative	Apologize
Abundance	Abolish	Agreement	Aggressive	Administrator	Apology
Abundant	Abominable	Aid#1	Aggravate	Admiration	Apprehensive
Accede	Abrasive	Alive	Agitate	Admirer	Ashamed
Accentuate	Abrupt	Alliance	Agitation	Admonish	Asleep
Accept	Abscond	Allied	Allegation	Adorn	Astray
Acceptable	Absence	Allow#1	Ambiguity	Adroit	Asunder
Acceptance	Absent-Minded	Ally#1	Ambivalent	Adroitly	Atrophy
Accessible	Absentee	Ameliorate	Amiss	Adulation	Average#1
Accession	Abuse	Ample	Antagonistic	Adult#1	Avert
Acclaim	Abyss	Appeal#3	Arbitrary	Advance#1	Avoid
Acclamation	Accident	Applause	Argue	Advantage	Avoidance
Accolade	Acrimonious	Appoint#1	Artificial	Advantageous	Awkward
Accommodate	Adulterate	Appreciable	Attack	Affinity	Awkwardness

Figure 3.1

Extensive form of imperfect information game between shareholders and management

The following figure is an extensive-form diagram of an imperfect information sequential game played between shareholders of a company and the manager of the company. Nature is also considered to be a player in the game and selects an action with some probability; this is the conventional game theoretic depiction of a random choice. Each cell represents a decision taken by one of the players: the shareholders, the manager or nature. The arrows depict the outcome of the decision. A dotted line represents an information set. At the end nodes, payoffs to the manager and shareholders are represented by values in parenthesis separated by a comma. S represents whether a lawsuit is chosen, F whether fraud is committed, and r whether a restatement occurs or not. A normal-form representation of this game with the same payoffs is presented in Chapter 3. See Appendix A for further details of the solution to this game.

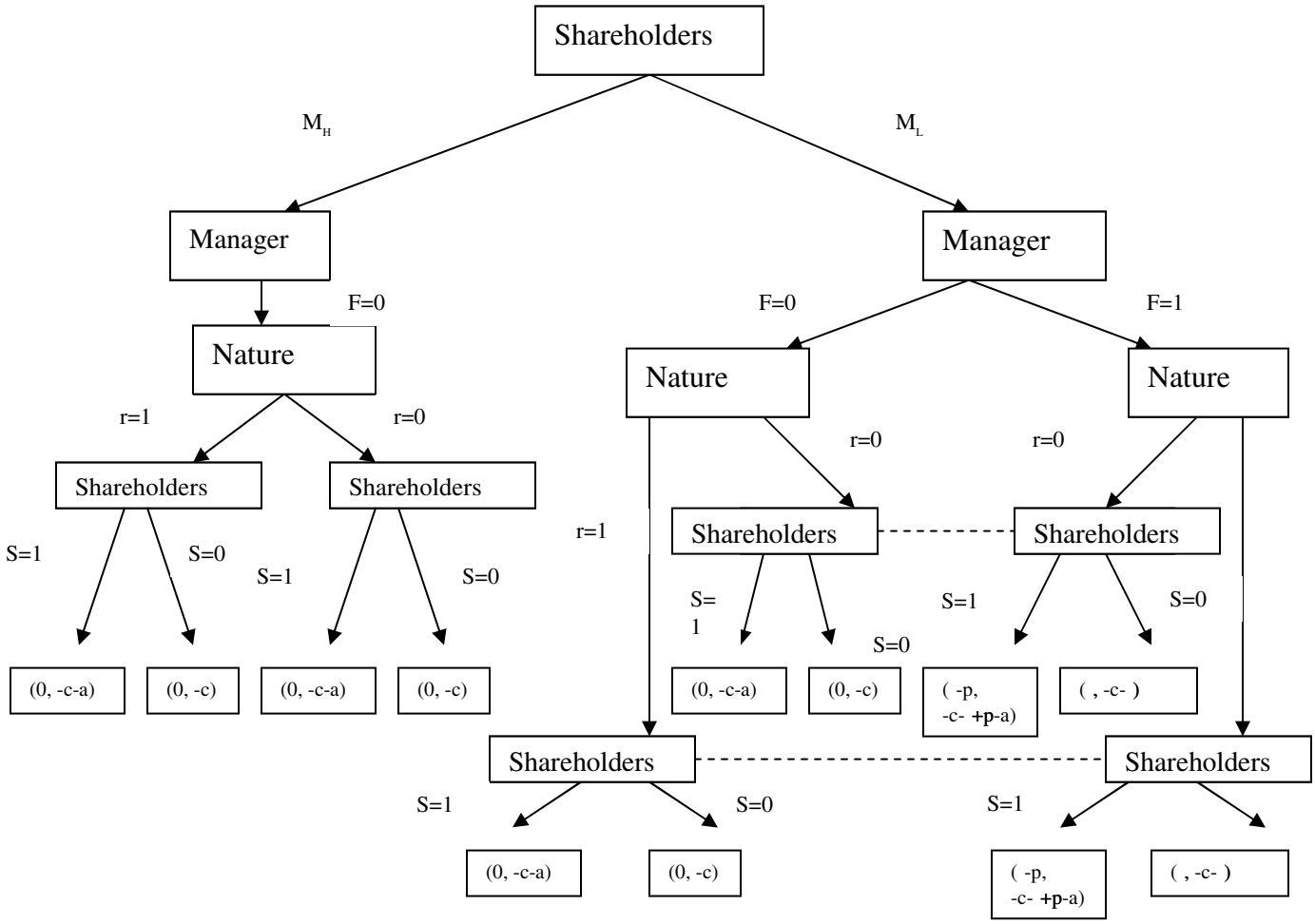


Figure 4.1

Securities class action lawsuits over time

This figure is a graphical depiction of the number of securities class action lawsuits filed in the post-reform act period. It covers the years 1996 to 2005 and includes all types of securities lawsuits during the period. All data is obtained from the Stanford Securities Class Action Clearinghouse, jointly maintained by Cornerstone Research. The large spike in 2001 is due to an unprecedented number of IPO allocation based lawsuits that were brought to court in that year. This graph includes securities class action lawsuits brought against private as well as public companies, as well as lawsuits brought against investment banks or traders for improper IPO practices. Not all the lawsuits depicted in this graph are in the sample used for regressions. Only lawsuits against public firms for which other data and control variables could be found are used in the regressions. IPO allocation lawsuits are excluded as far as possible.

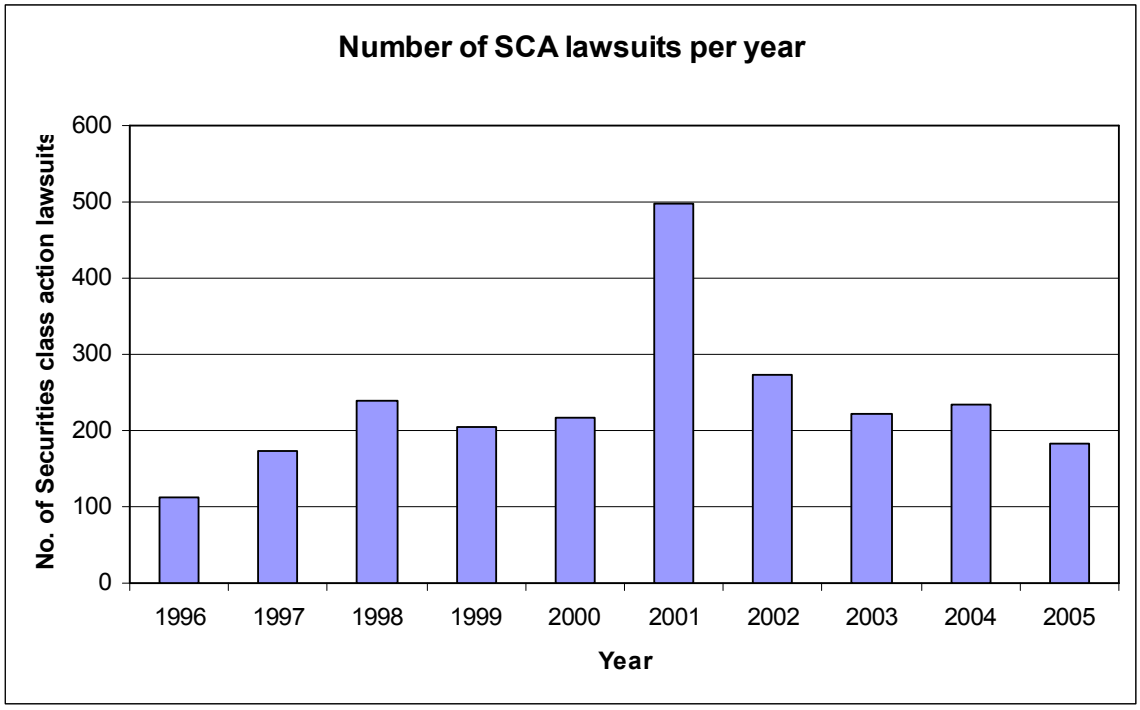
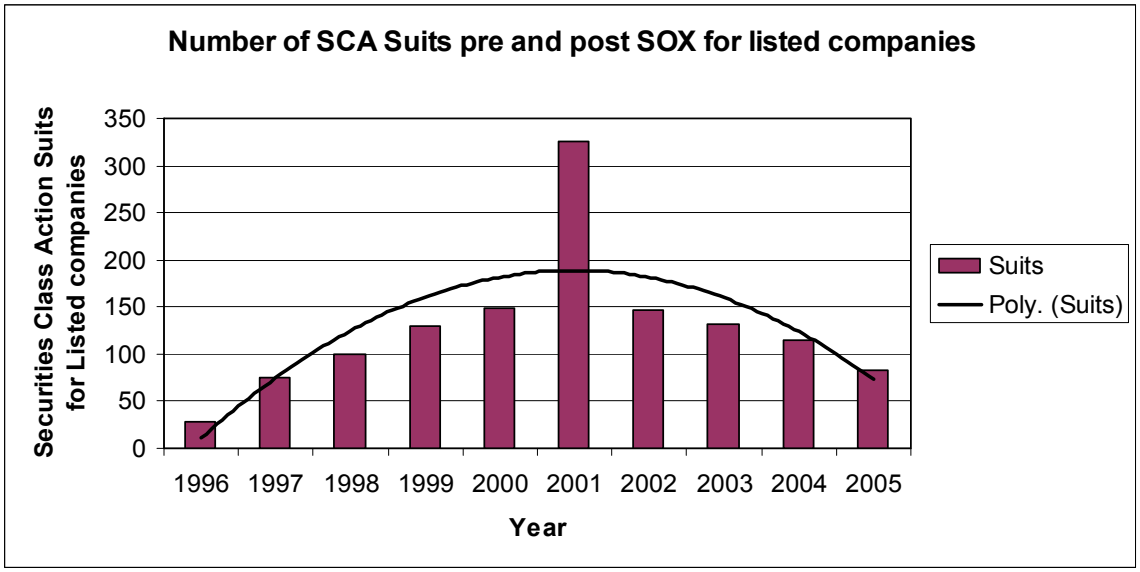


Figure 4.2

Change in number of securities class action lawsuits filed by public companies over time

This figure depicts all the securities class action lawsuits that have occurred in the years 1996 to 2005, in the post securities reform act period, for public companies only. The data includes all types of securities lawsuits during the period and is obtained from the Stanford Securities Class Action Clearinghouse, jointly maintained by Cornerstone Research. The large spike in 2001 is due to an unprecedented number of IPO allocation based lawsuits that were brought to court in that year. The growth in lawsuits is also fitted using a polynomial fit line. I also test to see if the number of lawsuits filed against public companies has decreased post Sarbanes Oxley. Since Sarbanes Oxley applies only to public companies, one may not expect to see a change in the lawsuits filed against private companies. Below the figure are the results of a Chow test for a structural break in the occurrence of lawsuits post-Sarbanes Oxley. I test for structural breaks at different dates in 2002, and all these tests indicate the presence of a structural break and a reduction in the number of lawsuits against public companies since the Sarbanes-Oxley Act was passed in 2002.



Chow test for structural breaks
(reduction in number of lawsuits)

Time	F Value	P- Value
Jan-02	30.74	<.0001
Jun-02	17.98	<.0001
Dec-02	19.9	<.0001

N=42338

Figure 5.1

Eigen values of largest principal components

The graph is a scree plot containing eigen values of the principal components extracted from the word categories as described in Chapter 5. The variance explained by each principal component is the eigen value of that component divided by the total of the eigen values. Numbers next to each point on the graph are indexes of the principal components. The names and constituents of the principal components are available in Table 5.7 and descriptive statistics are available in Table 5.8.

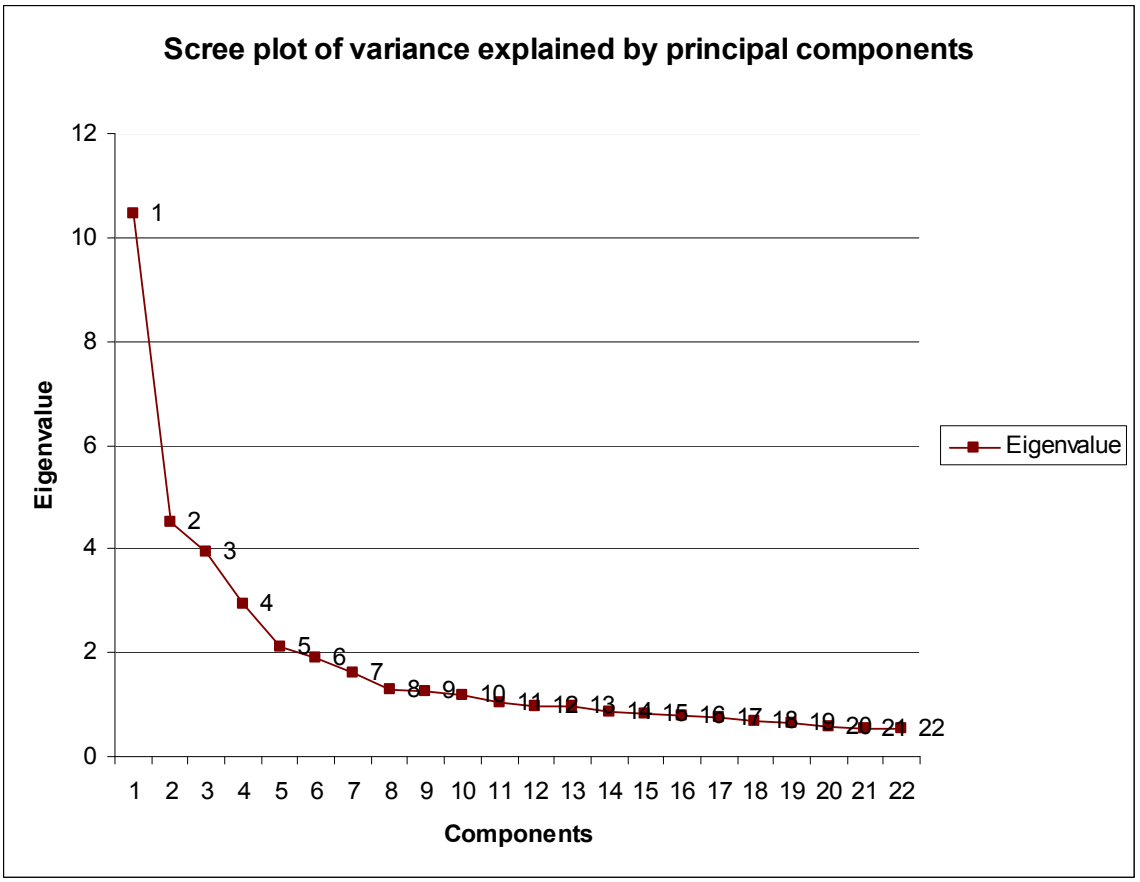


Table 4.1

Industries with the most securities class action lawsuits

This table lists the industries (defined by 4-digit SIC codes) which have the maximum number of securities class action lawsuits in the sample period. It covers the years 1996 to 2005 and includes all types of securities lawsuits during the period. All data is obtained from the Stanford Securities Class Action Clearinghouse, jointly maintained by Cornerstone Research. Industry definitions are obtained from the SIC industry descriptions.

Rank	Industry	4-digit SIC code	Number of lawsuits
1	Prepackaged Software	7372	35
2	Pharmaceuticals	2834	21
3	Computer programming,data processing	7370	12
4	Semiconductors	3674	10
5	Electric Services	4911	10
6	Electric and other services combined	4931	9
7	Computer communications equipment	3576	8
8	Telephone and telegraph apparatus	3661	8
9	Radio, TV broadcasting	3663	8
10	State commercial banks	6022	7

Table 4.2

Top ten allegations in securities class action lawsuits

This table lists the top ten allegations made by plaintiffs against management while filing a securities class action lawsuit. Lawsuits occurring in the years 1996 to 2005 are included in the sample. All data is obtained from the Stanford Securities Class Action Clearinghouse, jointly maintained by Cornerstone Research. The allegation terminology is as listed in the Stanford Securities Class Action Clearinghouse, for which data are obtained directly from lawsuit filings.

Rank	Allegation
1	Misleading/False Statements
2	Failure to disclose relevant information
3	Disclosure violation
4	Breach of fiduciary responsibility
5	Registration proxy statement
6	Insider trading
7	Revenue restatement
8	IPO
9	Improper accounting practice
10	Revenue recognition

Table 4.3

Ten largest settlements in securities class action lawsuits

This table describes the evidence that securities class action lawsuits can be a big financial burden to the company, even if they do not go to trial. A very large proportion of securities lawsuits never go to trial and are instead settled out of court. The table lists only settlements after the Reform Act was passed. Thus only settlements from lawsuits occurring in the years 1996 to 2005 are included in the sample. These numbers are as of February 13, 2006 and any additional settlements made after that date have not been included in these figures. The figure for all other large settlements given in the last row is the total of all other settlements whose amount is in excess of $100 million. All data is obtained from the Stanford Securities Class Action Clearinghouse, jointly maintained by Cornerstone Research.

Rank	Issuer	Settlement amount	Percentage of Valuation
1	Enron	$7,160.50	21.12%
2	WorldCom	$6,156.30	18.16%
3	Cendant	$3,528.00	10.41%
4	AOL Time Warner	$2,500.00	7.37%
5	Nortel Networks	$2,473.60	7.30%
6	Royal Ahold	$1,091.00	3.22%
7	IPO Allocation Litigation	$1,000.00	2.95%
8	McKesson HBOC	$960.00	2.83%
9	Lucent Technologies	$673.40	2.19%
10	Bristol-Myers Squibb	$574.00	1.69%
	All other large settlements	$7,786.50	22.97%

Table 4.4

Pearson correlation coefficients between lawsuits and monitoring variables

This table lists Pearson correlation coefficients between the different variables used in the study described in Chapter 4. Panel A contains the correlation of different proxy variables for institutional monitoring with the occurrence of a lawsuit. The occurrence of a lawsuit is a dummy variable that takes the value of 1 if a lawsuit has occurred in that firm-year. Monitoring variables are generally considered in the period one year prior to the year of the lawsuit. Panel B contains correlation coefficients between firm characteristics and the occurrence of a lawsuit. The P-Values in both Panel A and Panel B represent the statistical significance of the correlation coefficients.

Panel A

Variable	Correlation with lawsuits	P-Value	N
Blockholder existence	-0.0648	0.0009	4496
Institutional blockholder existence	-0.1254	<.0001	4496
Number of block holders	-0.0434	0.0794	4496
Proportion of independent directors	-0.0771	<.0001	2681
Total directors	-0.1675	<.0001	2681
CEO director	-0.0324	0.0924	4496
Number of institutional block holders	-0.1302	<.0001	4496
Total holdings by institutions	-0.2433	<.0001	4496
Herfindahl concentration of holdings	-0.0521	0.0075	2635

Panel B

Variable	Correlation with lawsuits	P-Value	N
Log(marketcap)	0.4451	<.0001	3842
Log(ROE)	-0.2089	<.0001	3712
CEO Salary	0.0875	<.0001	3605
CEO Option pay	0.1415	<.0001	3580
CEO Total pay	0.1170	<.0001	3580
Shares owned by CEO	0.0082	0.7579	1429
Dividend yield	-0.0201	0.2278	3586
Black-Scholes volatility of stock	0.0434	0.0124	3322
Tangible assets (PPE/Sales)	-0.0122	0.5067	2970
Total CEO pay/ Sales	0.0082	0.6264	3563

Table 4.5

Characteristics of restatements

The following table reports the results of logit regressions with restatement as the dependent variable. The restatement variable takes a value of 1 for firms that had a restatement of financial statements for that year and 0 otherwise. The firm characteristics include the natural logarithm of the firm's sales, compensation variables (total compensation and incentive-based pay of the CEO and the top five highest-paid executives of the company), board characteristics (the natural logarithm of the total number of directors on the board and the proportion of independent directors on the board) and the firm's dividend yield. The regression also includes year dummies. Incentive-based compensation is defined as the total of the Black-Scholes value of option grants and restricted stock grants, divided by the total pay. P-Values are given below each coefficient in italics. Compensation characteristics are obtained from Execucomp and all other firm data is from the Compustat database.

Independent variables:	(1)	(2)	(3)	(4)	(5)
Ln(Sales)	0.1444	0.1385	0.1127	0.1795	0.1666
	0.0061	*0.0095*	*0.0612*	*0.0011*	*0.0062*
Total Pay CEO		2.5149	0.4954		
		0.5215	*0.9235*		
Incentive Pay CEO	0.4943	0.4538	0.5353		
	0.079	*0.1165*	*0.0937*		
Total Pay Top5				9.4979	-2.1894
				0.4486	*0.8873*
Incentive Pay Top5				1.0394	0.9512
				0.0026	*0.0099*
Ln(Totaldirectors)			-0.1273		-0.1486
			0.452		*0.3689*
Prop. Indep. Directors			-0.5449		0.4013
			0.2572		*0.3194*
Dividend Yield					-0.5458
					0.8153
Year Dummies	Yes	Yes	Yes	Yes	Yes
N	10108	10108	5332	10260	6107

Table 4.6

Differences in monitoring levels between lawsuit and no-lawsuit subsamples

This table reports the results of univariate tests of differences between lawsuit firms and firms which have not experienced a lawsuit. Panel A compares mean values of monitoring variables and Panel B compares mean values of firm characteristics and CEO entrenchment variables, between two sets of firms. The first set of firms are those which have experienced a securities class action lawsuit in the sample period, and the second set of firms are control firms which have not experienced lawsuits. Blockholder existence is an indicator variable that takes a value of 1 if an institutional blockholder exists who owns 5% or more of the stock of the company. The last column reports P-Values from a Wilcoxon test comparing the mean values of the two sets of firms. Compensation data in Panel B is reported in thousands of dollars, as is Total Sales, Total Assets and Tangible Assets.

Wilcoxon tests for differences in monitoring between lawsuit and no-lawsuit firms

Panel A

		No lawsuit firms	Lawsuit firms	Wilcoxon test P-value
	Number of blockholders	2.0222	2.1973	0.0888
	Number of indep. directors	7.1676	5.9770	<.0001
	Total directors	10.4814	9.1593	<.0001
	Proportion of indep. directors	0.6816	0.6500	0.0001
	Blockholder existence	0.7312	0.6491	<.0001
	Number of analysts	10.1267	7.8827	<.0001
	Total holdings by institutions	0.5270	0.4294	<.0001
Observations		5251	1462	

Panel B

Wilcoxon tests for differences in firm characteristics between lawsuit and no-lawsuit firms

		No lawsuit firms	Lawsuit firms	Wilcoxon test P-value
Entrenchment	CEO salary	684.6332	871.3469	<.0001
	CEO bonus	840.2409	1408.2200	0.0785
	CEO total pay	5245.6200	17946.1200	<.0001
	CEO option pay	2486.9400	13769.1700	<.0001
	Top 5 total pay	8449.7800	24380.2100	0.0041
	Top 5 option pay	3970.3300	19250.1900	<.0001
	Dividends per share	0.5892	0.2831	<.0001
Firm characteristics	Sales	5940.4100	14118.6900	<.0001
	Employees	22.9979	40.3105	0.0024
	ROA	4.1170	2.3179	0.0004
	Assets	18081.0800	24288.0600	<.0001
Complexity	PPE (Tangible assets)	403.0990	462.2820	<.0001
	Stdev. of analyst forecasts	0.4208	0.3049	<.0001
	Black-Scholes Volatility	0.3463	0.3871	<.0001

Table 4.7

Logistic regressions of the occurrence of a lawsuit on monitoring variables

The following table reports the results of logit regressions with lawsuits as the dependent variable. The lawsuit variable takes a value of 1 for firms that had a securities class action lawsuit in the year and 0 otherwise. This is a panel data regression with the standard errors adjusted for clustering, or multiple observations per firm. The independent variables include the following controls: log of marketcap of the firm as a control for size, log of return on equity both current and lagged, the previous fiscal year's stock returns and the occurrence of a restatement. I also control for discretionary accruals which is considered to be a sign of earnings manipulation, and may thus trigger a lawsuit. Monitoring variables include institutional blockholder monitoring (a dummy value that takes a value of '1' if an institutional blockholder exists), board member monitoring (the percentage of independent directors on the board), and analyst monitoring (number of analysts and the standard deviation of forecasts by analysts). The blockholder dummy is also interacted with (1- discretionary accruals). This interaction term takes a high value if there exists at least one institutional blockholder, and the firm does not show signs of earnings manipulation. The regression also includes year and Fama-French industry dummies. P-Values and marginal effects (slope of the regression line for an average value of the covariates) are given adjacent to the coefficient estimates.

Dep. Variable: Suit	Coefficient	P-Value	Marginal effect	Coefficient	P-Value	Marginal effect	Coefficient	P-Value	Marginal effect	Coefficient	P-Value	Marginal effect
Intercept	-0.77	*0.3381*		-7.7871	*<.0001*		-1.457	*0.1078*		-5.4495	*<.0001*	
Restatement	0.9068	*0.0001*		1.383	*<.0001*		0.4251	*0.0459*		0.8312	*0.0046*	
Log(Size)	0.0325	*0.3512*	*5.1283*	0.3573	*<.0001*	*5.1171*	0.0107	*0.8075*	*5.1728*	0.3085	*<.0001*	*3.5788*
Log(ROE)	-1.0594	*<.0001*	*0.0145*	-2.3141	*0.0011*	*0.0144*	-1.3961	*<.0001*	*0.0146*	-2.5525	*<.0001*	*0.0101*
Fisc year return	0.104	*0.2933*	*0.011*	0.7258	*0.0002*	*0.0109*	0.0426	*0.6485*	*0.0111*	0.1044	*0.2915*	*0.0076*
Disc. Accruals										-0.0001	*0.2244*	
Blockholder dummy	-0.4875	*<.0001*								-0.4151	*0.0186*	
Blockholder x (1-Disc.acc.)										-0.0002	*0.0287*	
PC independent directors				-1.1829	*0.0021*	*0.1641*				-1.6764	*<.0001*	*0.1642*
Number of analysts							-0.0058	*0.5227*	*2.3376*			
Stdev of analyst forecasts							0.7799	*0.1158*	*0.0953*	-2.5492	*0.1288*	*0.0943*
Year dummies	Yes			Yes			Yes			Yes		

Fama-French industry dummies	Yes	Yes	Yes	Yes
N	6731	6731	6731	6731

Table 4.8

Tests for endogeneity between lawsuit and monitoring variables

This table contains the results of tests for endogeneity of the blockholder variables in the lawsuit regression. The test examines if the error term is correlated with this institutional holding variable when lawsuits are regressed on either a blockholder dummy variable, or the proportion of shares held by institutions. Results of the test show that when either of these proxies is used to measure monitoring by institutions, an OLS does not result in unbiased estimates of the coefficient and a Two-Stage Least Squares regression method is preferable to OLS.

Test for endogeneity: Choosing between OLS and 2SLS regressions

Possibly endogenous variables	Comparing	To	DF	Test Statistic	P-Value (- squared)
Blockholder existence and lawsuits	OLS	2SLS	2	4.871	<0.0001
Institutional holdings and lawsuits	OLS	2SLS	2	6.842	<0.0001
N				21973	

Table 4.9

Two-stage least squares estimation of the relationship between monitoring and lawsuits

This table shows the results of a Two-Stage Least Squares regression in which membership in the S&P index is used as an instrument for institutional monitoring. This instrument is used to perform logistic regressions of the occurrence of a lawsuit on the institutional monitoring variable. First and second stage regression results are given in Panel A and B respectively. The three first stage regressions have as a dependent variable the proportion of shares held by institutions, the existence of a blockholder, and the number of block holders respectively. In all these three regressions, the dependent variable is S&P index membership and the coefficients in the table are the coefficients on this variable. P-Values and the adjusted R-squared are also displayed. In the second stage regressions, the occurrence of a lawsuit is the dependent variable and the various independent variables are listed in the table. These are the same as used in Table 4.6 except that the first stage predicted values are substituted for the institutional monitoring variables: Proportion of institutional holdings, Blockholder existence and number of block holders. P-Values are given along with the coefficients.

Panel A: First stage regression using S&P instrument

Dep.Variable	Coefficient	P-Value	R-Squared
Prop. Inst. Holding	0.2817	<.0001	0.0001
Blockholder	0.0045	0.028	0.0002
# Block holders	0.9466	<.0001	0.0163

Panel B: Second stage regression using S&P instrument

Dep.Var. : Suit	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Intercept	-0.2105	<.0001	-0.1305	<.0001	-0.1803	<.0001
Restatement	0.1721	<.0001	0.1570	<.0001	0.1570	<.0001
Ln(Size)	0.0131	<.0001	0.0131	<.0001	0.0131	<.0001
Ln(ROE)	-0.0082	0.0074	-0.0082	0.0078	-0.0082	0.0103
Lag Ln(ROE)	-0.0048	0.1356	-0.0037	0.244	-0.0037	0.2615
Fisc.yr. return	-0.0358	<.0001	-0.0349	<.0001	-0.0349	<.0001
Prop. Inst. Holding	-0.0782	<.0001				
Blockholder			-0.1511	<.0001		
# Block holders					-0.0239	<.0001
Year dummies		Yes		Yes		Yes
FF industry dummies		Yes		Yes		Yes
N		22072		22072		22072

Table 4.10

Marginal effects of a restatement on the probability of a lawsuit

The following table reports descriptive statistics for marginal changes in the value of the 'suit' variable when a restatement occurs, for different types of firms. The first two rows report average marginal effect on the probability of a lawsuit for firms with and without institutional blockholders. An institutional blockholder is an institutional investor holding 5% or more of the shares of the company. The rows also report the minimum and maximum marginal effects as well as the number of observations. The next two rows report the marginal effects of a restatement on the probability of a lawsuit for two groups of firms according to their proportion of institutional holdings. This is calculated as the percentage of institutional shareholdings divided by the total shares outstanding.

Subsamples	Mean	Minimum	Maximum	N
Firms with institutional blockholder	0.0800	0.0058	0.7621	3623
Firms without institutional blockholder	0.1441	0.0125	0.6552	3167
Firms with above-average proportion of institutional holdings	0.0605	0.0000	0.4873	3538
Firms with below-average proportion of institutional holdings	0.0944	0.0145	0.7060	7668

Table 4.11

Changes in institutional holdings of shares in lawsuit firms over time

The following table reports results from an event-time regression in which the proportion of shares held by investors is the dependent variable. All firms included in the sample are firms which have had securities class action lawsuits during the sample period. Year 0 refers to the year of the lawsuit, years -1, -2 etc. refer to years prior to the lawsuit and years +1,+2 etc refer to years after the lawsuit has been filed. Other than using year and Fama-French industry dummies to control for these effects, I also control for some common determinants of institutional holdings. These include size (logged market capitalization), logged book to market, leverage, stock returns during the prior year, and share turnover or liquidity. In addition, I control for the occurrence of a restatement and for earnings manipulation using discretionary accruals.

	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Intercept	0.4223	<.0001	-0.8902	<.0001	-0.6484	<.0001
Year -5	0.0271	0.2374	0.0125	0.568	0.0407	0.5658
Year -4	0.0488	0.0282	0.0301	0.1349	-0.1055	0.1386
Year -3	0.0524	0.0179	0.0261	0.202	-0.0752	0.1762
Year -2	0.0189	0.3779	0.0333	0.0793	-0.0997	0.0522
Year -1	-0.0066	0.7483	0.0082	0.6511	-0.1455	0.0085
Year 0	-0.0568	0.0031	-0.0320	0.0493	-0.2155	0.0005
Year + 1	-0.0590	0.0007	-0.0435	0.006	-0.0370	0.6311
Year +2	-0.0417	0.01	-0.0185	0.2451	0.0185	0.8099
Year +3	-0.0457	0.0025	-0.0026	0.8649	-0.1190	0.3296
Ln(Marketcap)			0.0666	<.0001	0.0606	<.0001
Ln(B/M)			0.0347	<.0001	0.0528	0.0002
Fisc. Year returns			0.0287	<.0001	0.0379	0.0013
Leverage			0.0001	0.407	0.0036	0.0966
Share turnover			0.0846	0.0036	0.0447	0.1596
Restatement					-0.0027	0.9257
Disc. Accruals					0.0000	0.0008
Calendar year dummies		Yes		Yes		Yes
FF industry dummies		Yes		Yes		Yes
N (clusters)		791		781		696

Table 4.12

Does monitoring directly deter fraud? Logistic regressions using matched sample firms

This table reports the results of a matched-sample (pairwise) regression using 70 firms which were investigated by the SEC for alleged fraud (the sample firms). These sample firms are then matched with 70 control firms, which are firms in the same SIC 4-digit industry code closest in size and past profitability to the sample firms in the years before the alleged fraud took place. I test whether strong monitoring deters fraud by regressing the occurrence of fraud on various monitoring variables and controls. Monitoring variables include monitoring by institutional investors (proportion of shares held by institutions), monitoring by boards (total number of directors on the board of a company, and the proportion of inside directors on the board), and monitoring by analysts (number of analyst forecasts). I also control for size (log of market capitalization of the company), profitability (logged ROE), leverage and discretionary accruals.

	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Discretionary accruals	0.0003	0.7164	0.0005	0.7912	0.0003	0.6377	0.0002	0.6041	0.0001	0.8482
Ln (Marketcap)	2.6162	0.1492	3.0739	0.2445	0.7067	0.0335	0.7240	0.0755	1.0344	0.1054
Ln(ROE)	-13.6595	0.2615	-15.6471	0.3380	-0.3917	0.5172	-0.2950	0.6236	-0.6928	0.5102
Leverage	0.6756	0.5399	1.1868	0.4931	0.3321	0.2679	0.3086	0.3353	-0.0052	0.9920
PC indep directors	1.0513	0.7663	1.5184	0.7096			-0.7232	0.7162		
Total directors					0.3231	0.4347				
Block holders					0.0637	0.5044				
Institutional holdings										
# analysts									0.0589	0.2421
N		140		140		140		140		140

Table 5.1

Correlation coefficients between sets of LIWC word categories

This table reports Pearson correlation coefficients for different word categories available in the content analysis package LIWC (Linguistic Inquiry and Word Count) using the output from the entire dataset of 10-K reports filed with the SEC in the years 1993-2005. P-Values for the correlation coefficients are reported below the coefficients. WPS is Words per sentence. All category outputs barring Wordcount and Words per sentence are standardized by dividing them by the total wordcount.

LIWC word categories	Wordcount	WPS	Pronouns	Numbers	Tentative	Certain	Past	Present	Future
Wordcount	1	0.4766	0.0357	-0.4457	0.6258	0.3046	-0.3616	-0.2882	0.5426
		<.0001	0.0599	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
WPS	0.4765	1	-0.1018	-0.2030	0.5355	0.2310	-0.2562	-0.3738	0.3913
	<.0001		0.0195	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Pronouns	0.0357	-0.1018	1	-0.1098	0.0084	0.0025	-0.1615	0.2829	0.0248
	<.0001	<.0001		<.0001	0.0653	0.5823	<.0001	<.0001	<.0001
Numbers	-0.4457	-0.2030	-0.1098	1	-0.7373	-0.4809	0.3303	-0.0491	-0.7443
	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
Tentative	0.6257	0.5355	0.0084	-0.7373	1	0.5131	-0.4588	-0.2797	0.8683
	<.0001	<.0001	0.0653	<.0001		<.0001	<.0001	<.0001	<.0001
Certain	0.3046	0.2310	0.0025	-0.4809	0.5131	1	-0.2305	-0.1190	0.3985
	<.0001	<.0001	0.5823	<.0001	<.0001		<.0001	<.0001	<.0001
Past	-0.3615	-0.2562	-0.1615	0.3303	-0.4588	-0.2305	1	0.0666	-0.4325
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001
Present	-0.2881	-0.3738	0.2829	-0.0491	-0.2798	-0.1191	0.0666	1	-0.1756
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001
Future	0.5426	0.3914	0.0248	-0.7443	0.8684	0.3985	-0.4326	-0.1756	1
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.002	

Table 5.2

Correlation coefficients between sets of GI word categories

This table reports the Karl Pearson correlation coefficients for different word categories available in the content analysis package General Inquirer (GI), using the output from the entire dataset of 10-K reports filed with the SEC in the years 1993-2005. P-Values for the correlation coefficients are reported below the coefficients. WPS is Words per sentence. All category outputs barring the wordcount and words per sentence are standardized by dividing them by the total wordcount. Two different measures of positive and negative words are used which use inputs from different dictionaries.

	Wordcount	Positive	Negative	Positive 2	Negative 2	Leftovers	Strong	Weak	Overst.	Underst.
Wordcount	1	0.1075	0.0568	0.2042	-0.0767	-0.2319	-0.2837	-0.1733	0.6000	-0.1833
		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Positive	0.1075	1	-0.0042	0.9050	-0.0120	-0.2900	0.1732	0.1189	0.1349	0.0824
	<.0001		0.3547	<.0001	0.0087	<.0001	<.0001	<.0001	<.0001	<.0001
Negative	0.0568	-0.0042	1	-0.0492	0.9413	-0.2000	0.0766	0.5204	0.0610	0.2740
	<.0001	0.3547		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Positive 2	0.2042	0.9050	-0.0492	1	-0.0784	-0.3516	0.1435	-0.0373	0.2524	0.0365
	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Negative 2	-0.0767	-0.0120	0.9413	-0.0784	1	-0.1241	0.1668	0.5566	-0.1171	0.3108
	<.0001	0.0087	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
Leftovers	-0.2319	-0.2900	-0.2000	-0.3516	-0.1241	1	-0.2722	0.0142	-0.4855	-0.0585
	<.0001	<.0001	<.0001	<.0001	<.0001		0.0019	0.0019	<.0001	<.0001
Strong	-0.2837	0.1732	0.0766	0.1435	0.1668	-0.2722	1	0.1069	-0.2753	0.1963
	<.0001	<.0001	<.0001	<.0001	<.0001	0.0019		<.0001	<.0001	<.0001
Weak	-0.1733	0.1189	0.5204	-0.0373	0.5566	0.0142	0.1069	1	-0.2201	0.2665
	<.0001	<.0001	<.0001	<.0001	<.0001	0.0019	<.0001		<.0001	<.0001
Overstatement	0.6000	0.1349	0.0610	0.2524	-0.1171	-0.4855	-0.2753	-0.2201	1	-0.1891
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001
Understatement	-0.1833	0.0824	0.2740	0.0365	0.3108	-0.0585	0.1963	0.2665	-0.1891	1
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	

Table 5.3

Significant differences in word usage between lawsuit and no-lawsuit subsamples

Mean values of various word categories for 10-K reports filed by two subsamples: lawsuit firms and firms that have not experienced a lawsuit are reported here. WC is wordcount and WPS is words per sentence. Figures for word categories are reported in terms of a percentage of the total words contained in the document. The P-Values in the table are associated with Wilcoxon tests for differences in the means of the word categories values between the two subsamples and are used to indicate whether the two subsample means are significantly different from each other. Panel A lists the subsample means for word categories in the content analysis package Linguistic Inquiry and Word Count (LIWC). Panel B lists the subsample mean values for word categories in the General Inquirer (GI) content analysis package.

Panel A: Differences in LIWC word categories

Word categories	No lawsuit sample	Lawsuit firm-years	P-Value
WC	53766	63116	0.0006
WPS	35.5959	35.8819	0.8032
Pronoun	1.1030	1.0990	0.6231
Negate	0.6457	0.6388	0.8066
Assent	0.3931	0.4062	0.0783
Number	8.0134	7.2921	0.0000
Tentat	2.1594	2.3832	0.0033
Certain	0.5141	0.5270	0.0315
Past	1.4400	1.3948	0.0000
Present	2.8840	2.8777	0.6210
Future	1.4812	1.5639	0.0459
Achieve	0.8038	0.8077	0.4472
Allpct	33.7555	32.9462	0.9257
Negemo	0.3825	0.3919	0.1238
Posemo	2.6835	2.7074	0.6290
Optim	0.9341	0.9428	0.9931
Posfeel	0.3453	0.3551	0.1915

Panel B: Differences in GI word categories

Word categories	No lawsuit sample	Lawsuit firm-years	P-Value
Wordcount	50930	60097	0.0006
Positive	0.0613	0.0616	0.8427
Negative	0.0214	0.0216	0.4412
Positive 2	0.0514	0.0517	0.6828
Negative 2	0.0171	0.0171	0.9977
Leftovers	0.1207	0.1195	0.7930
Strong	0.1301	0.1312	0.1412
Weak	0.0159	0.0155	0.0269
Overstatement	0.0415	0.0431	0.0084
Understatement	0.0080	0.0080	0.8590

Table 5.4

Significant differences in mean values of principal components between lawsuit and no-lawsuit subsamples

Mean values of various principal components based on word categories in GI and LIWC, calculated for 10-K reports filed by two subsamples: lawsuit firms and firms which have not experienced lawsuits, are reported here. The P-Values given in the table are associated with Wilcoxon tests for differences in the means of the variable between the two subsamples and are used to indicate whether the two subsample means are significantly different from each other. More details about these principal componetns and their constituents are in Table 5.6.

Principal Component	No lawsuit sample	Lawsuit firm-years	P-Value
PC1 Wordcount	4704	11284	0.0028
PC2 Punctuation	0.9922	1.5602	0.0676
PC3 Wordiness	-0.1000	-0.7194	0.0161
PC4 Informativeness	0.4297	-0.0898	0.0047
PC5 Intangibles	-0.1549	0.1510	0.0004
PC6 Factual references	0.3042	0.1783	0.0023
PC7 Predictive words	0.2447	0.1723	0.0265
PC8 Positive	-0.0552	-0.1060	0.4410
PC9 Small words and numbers	-0.0839	-0.2008	0.0003

Table 5.5

Unconditional logistic regressions of lawsuits on word categories

The table reports the results of unconditional logistic regressions in which the dependent variable is a binary-valued lawsuit variable that takes the value of 1 if a lawsuit has occurred for that firm-year observation and 0 otherwise. Results are reported for four different regression specifications. For each regression, the coefficient, P-Value and marginal effect are all reported. The regressions include word categories from the content analysis packages Linguistic Inquiry and Word Count (LIWC) and the General Inquirer (GI). The regressions also include controls for size and profitability (logged market capitalization and logged ROE respectively). Controls for future and past profitability and stock returns are included as possible determinants of legal action. Since earlier tests have shown that strong monitoring is negatively related to litigation, I also include controls for the existence of a blockholder and the percentage of independent directors on the board. LIWC and GI word categories are indexed as such. LIWC_WC is a control for the total number of words in the document. Other LIWC word categories used are negative emotions, positive emotions, pronouns, numbers, tentative words, past, present and future words. GI word categories in the regressions are positive words, negative words, strong and weak words, overstatement and understatement.

Dep. Var: Suit	Coeff.	*P-Value*	*M.Eff.*	Coeff.	*P-Value*	*M.Eff.*	Coeff.	*P-Value*	*M.Eff.*	Coeff.	*P-Value*	*M.Eff.*
		I			*II*			*III*			*IV*	
Intercept	-0.2787	*<.0001*		-0.3	*<.0001*		-0.3213	*<.0001*		-0.3058	*<.0001*	
Ln(Marketcap)	0.0149	*<.0001*	*0.3824*	0.0146	*<.0001*	*0.0074*	0.0146	*<.0001*	*-0.0535*	0.0146	*<.0001*	*-0.0103*
Ln(ROE)	-0.0018	*0.8705*	*-0.0001*	-0.0018	*0.8714*	*0.0000*	-0.0032	*0.7732*	*0.0000*	-0.0019	*0.8647*	*0.0000*
Restatement	0.2361	*<.0001*	*0.0000*	0.2345	*<.0001*	*0.0000*	0.2356	*<.0001*	*0.0000*	0.2347	*<.0001*	*0.0000*
Future Ln(ROE)	-0.0309	*0.0034*	*-0.0001*	-0.0302	*0.0042*	*0.0000*	-0.031	*0.0034*	*0.0000*	-0.0295	*0.0049*	*0.0000*
Lag Ln(ROE)	0.0075	*0.3426*	*0.0002*	0.0078	*0.3246*	*0.0000*	0.0063	*0.4304*	*0.0000*	0.0072	*0.3614*	*0.0000*
Fisc.yr. returns	-0.0461	*<.0001*	*-0.0016*	-0.0459	*<.0001*	*0.0000*	-0.0458	*<.0001*	*0.0002*	-0.0453	*<.0001*	*0.0000*
Blockholder	0.0095	*0.1148*	*0.0137*	0.0086	*0.1536*	*0.0003*	0.0088	*0.1394*	*-0.0019*	0.0077	*0.1942*	*-0.0004*
PC Indep Dirs	0.0205	*0.0749*	*0.0125*	0.0232	*0.0415*	*0.0002*	0.021	*0.0669*	*-0.0017*	0.0251	*0.0307*	*-0.0003*
LIWC_WC	0.0075	*0.0225*	*0.0011*	0.007	*0.0564*	*0.0000*				0.0076	*0.0358*	*0.0000*
LIWC_Pronoun										0.0049	*0.1696*	*0.0000*
LIWC_Negemo	0.0045	*0.0423*	*0.0013*	0.0038	*0.0909*	*0.0000*						

Table 5.5, cont.

LIWC_Posemo	-0.0027	*0.4906*	*-0.0003*	-0.0017	*0.5641*	*0.0000*						
LIWC_Number				-0.012	*0.0012*	*0.0000*				-0.0118	*0.0018*	*0.0000*
LIWC_Tentat				-0.0083	*0.1202*	*0.0000*				-0.0041	*0.5413*	*0.0000*
LIWC_Past				-0.0061	*0.0205*	*0.0000*				-0.0044	*0.1073*	*0.0000*
LIWC_Present				0.0013	*0.5912*	*0.0000*				0.0002	*0.9433*	*0.0000*
Dep. Var: Suit	Coeff.	*P-Value*	*M.Eff.*	Coeff.	*P-Value*	*M.Eff.*	Coeff.	*P-Value*	*M.Eff.*	Coeff.	*P-Value*	*M.Eff.*
		I			*II*			*III*			*IV*	
LIWC_Future				-0.0057	*0.138*	*0.0000*				-0.0053	*0.1935*	*0.0000*
GI_Wordcount							0.0001	*0.0052*	*-118.834*			
GI_Positiv							0.002	*0.4426*	*0.0000*	0.0013	*0.6316*	*0.0000*
GI_Negativ							0.0052	*0.0445*	*0.0002*	0.0047	*0.0648*	*0.0000*
GI_Strong							0.006	*0.0222*	*0.0000*	0.0053	*0.0468*	*0.0000*
GI_Weak							-0.0054	*0.0279*	*0.0002*	-0.0042	*0.0701*	*0.0000*
GI_Ovrst							-0.002	*0.5219*	*0.0000*	-0.0021	*0.6254*	*0.0000*
GI_Undrst							1.3057	*0.3103*	*0.0000*	0.7505	*0.5712*	*0.0000*

Table 5.6

Conditional logistic regressions of lawsuits on word categories

Regressions in Table 5.5 are re-estimated using a Chamberlain conditional logit approach in order to control for multiple observations through time for each firm in the sample. The dependent variable is a binary-valued lawsuit variable that takes the value of 1 if a lawsuit has occurred for that firm-year observation and 0 otherwise. Results are reported for four different regression specifications. The regressions include word categories from the content analysis packages Linguistic Inquiry and Word Count (LIWC) and the General Inquirer (GI). LIWC and GI word categories are indexed as such. The regression includes controls for size and profitability (logged market capitalization and logged ROE respectively). Controls for future and past profitability and stock returns are also included as possible determinants of legal action. Since earlier tests have shown that strong monitoring is negatively related to litigation, I also include controls for the existence of a blockholder and the percentage of independent directors on the board. LIWC_WC is a control for the total number of words in the document and LIWC_WPS is the number of words per sentence in the document. Other LIWC word categories used are negative emotions, positive emotions, pronouns, numbers, tentative words, certain words, past, present and future words. GI word categories in the regressions are positive words, negative words, strong and weak words, overstatement and understatement.

Dep.Var: Suit	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Ln(Size)	0.9218	<.0001	0.9161	<.0001	0.9248	<.0001	0.9322	<.0001
Ln(ROE)	0.2619	0.5507	0.1801	0.6795	0.2111	0.6116	0.124	0.7781
Restatement	1.5449	0.008	1.5006	0.0125	1.4442	0.0137	1.4691	0.0139
Future Ln(ROE)	-0.0569	0.832	-0.0795	0.767	-0.0332	0.9019	-0.0618	0.8209
Lag Ln(ROE)	1.6608	0.0336	1.5988	0.0393	1.6055	0.0365	1.612	0.0408
F. yr. return	-0.8863	<.0001	-0.8787	<.0001	-0.8626	<.0001	-0.9062	<.0001
Blockholder	-0.2331	0.4334	-0.1765	0.5523	-0.2787	0.3441	-0.1794	0.55
Pc indep. Directors	-1.4298	0.1441	-1.4296	0.1487	-1.3256	0.1741	-1.4262	0.1534
LIWC_WC	0.1324	0.0224	0.1156	0.0323			0.1177	0.0223
LIWC_WPS	-0.0942	0.5323	-0.0412	0.8448			-0.042	0.8449
LIWC_Negemo	-0.0463	0.8046	-0.163	0.4259			-0.0253	0.9202
LIWC_Posemo	-0.0261	0.9095	-0.2536	0.1838			-0.221	0.3228
LIWC_Pronoun							-0.2184	0.2323
LIWC_Number			-0.0547	0.0285			-0.074	0.0172
LIWC_Tentat			0.0402	0.9009			0.1939	0.646
LIWC_Certain			0.1885	0.2626			0.1618	0.3656
LIWC_Past			-0.0729	0.0631			-0.0613	0.0399
LIWC_Present			0.1718	0.3024			0.1629	0.3491
LIWC_Future			-0.2564	0.0269			-0.2991	0.0219
GI_Wordcount					2.77E-06	0.3295		
GI_Positive					0.1117	0.4627	-0.0393	0.8534
GI_Negative					-0.0168	0.9187	-0.0808	0.6882
GI_Strong					0.2013	0.0306	0.1461	0.0493
GI_Weak					-0.1974	0.026	-0.1016	0.0582
GI_Ovrst					0.0435	0.8339	-0.0459	0.8898
GI_Undrst					15.2869	0.8762	-0.0017	0.998

Table 5.7

Description of principal components

The following are the descriptions of the main principal components used in the analysis, along with their constitutent word categories. The names of the components are descriptive and have been assigned based on the categories that comprise them. Figure 5.1 is a scree plot that details the eigen values of the principal components.

Principal Components	Name of component	Constituent categories
1	Document length	Wordcount from GI and LIWC
2	Punctuation	Punctuation
3	Wordiness	Words per sentence
4	Informativeness	Unique words and Numbers, (-) Tentative words
5	Intangibles	Unique words, - Numbers
6	Factual references	Articles, (-) Pronouns
7	Predictive words	Tentative words, Future words, Numbers
8	Positive words	Positive emotional words, Optimistic words, Affecting words
9	Small words and numbers	Prepositions, Pronouns, Numbers

Table 5.8

Summary statistics related to principal components

Summary statistics regarding each of the 11 principal components with the highest eigen values are listed here. These principal components were extracted from the scores of the 10-K documents along various dimensions of the LIWC and GI word categories. The selected 11 principal components together explain more than 95 percent of the variation in the 10-K documents.

Component	Min	Max	Mean	Standard Deviation	Eigen value	Cum. variance explained
Component 1	-15.9494	15.1663	0.0367	3.3458	10.7261	0.3462
Component 2	-25.0438	10.4475	-0.2549	2.0648	4.2034	0.4818
Component 3	-9.7061	18.8293	-0.0352	2.0025	3.7645	0.6033
Component 4	-8.9640	16.9181	-0.0651	1.7396	2.7296	0.6914
Component 5	-5.2992	20.6272	0.0788	1.4808	1.9355	0.7539
Component 6	-21.6371	5.0476	0.1396	1.3624	1.5919	0.8052
Component 7	-7.4798	14.1618	-0.0654	1.2631	1.3319	0.8482
Component 8	-7.0508	36.4893	-0.0530	1.0737	0.9565	0.8791
Component 9	-26.0591	35.7533	0.0838	1.0417	0.8872	0.9077
Component 10	-10.8770	18.6697	-0.0490	1.0979	0.8408	0.9349
Component 11	-22.4828	67.4228	0.0008	1.0348	0.6402	0.9555

Table 5.9

Conditional logistic regressions of lawsuits on principal components

Results from a Chamberlain conditional logit model in which the occurrence of a lawsuit is regressed on determinants of a lawsuit and principal components of the characteristics of an annual report, are reported here. I regress this lawsuit variable on all the different principal components 1 through 12, and controls for the determinants of a lawsuit and monitoring variables. Controls include logged size, present, past and future logged profitability, and the occurrence of a restatement. Controls for monitoring by institutions, blockholders and independent directors on the board of directors are also included. Blockholder is an indicator variable that takes a value of 1 if there exists at least one institutional blockholder holding 5% or more of the shares. The institutional holdings variable represents the proportion of shares held by institutions and PC Indep. Dirs. represents the percentage of independent directors on the board of the company.

Dep. Var: Suit	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Ln(Size)	0.5199	<.0001	0.5491	<.0001	0.7273	0.0004	0.7534	<.0001
Ln(ROE)	-0.1706	0.3694	-0.116	0.5377	-0.2087	0.5835	-0.0387	0.9203
Restatement	2.1307	<.0001	1.8699	<.0001	2.2873	0.0039	1.653	0.0042
Future Ln(ROE)	-0.3311	0.0102	-0.3239	0.0104	-0.3732	0.0799	-0.3457	0.1028
Lag Ln(ROE)	0.7186	0.004	0.6751	0.0044	1.4328	0.035	1.6294	0.0155
Inst. Holdings	0.4848	0.3463			-0.0085	0.9934		
Blockholder			0.2835	0.1191			0.2708	0.3323
PC Indep. Dirs.					-1.3232	0.202	-0.892	0.3138
Component1			0.0000	0.0749	0.0000	0.2265	0.0000	0.1213
Component2			0.0226	0.0054	0.0089	0.4871	0.0134	0.2385
Component3			-0.0114	0.4742	-0.0091	0.7011	-0.0214	0.3950
Component4	-0.0295	0.0334	-0.0929	0.0314	-0.0121	0.0851	-0.0488	0.0446
Component5			-0.0580	0.3200	-0.0951	0.3036	-0.1667	0.0579
Component6			0.1205	0.2180	-0.0444	0.8039	0.0884	0.5804
Component7			-0.2700	0.0243	-0.2565	0.0253	-0.2918	0.1073
Component8			0.0178	0.8795	0.1463	0.3709	0.1045	0.4893
Component9			-0.0452	0.7317	-0.0657	0.7588	0.0612	0.7611
Component10			0.0542	0.7648	0.1006	0.7236	-0.0310	0.9053
Component11			-0.3831	0.0537	-0.6683	0.0333	-0.3779	0.1653
Component12			0.1336	0.5978	0.6549	0.1106	0.5450	0.1351
N	24266		24266		24266		24266	

Table 5.10

Correlation coefficients of principal components and S&P ranks with lawsuits

The table reports Pearson correlation coefficients and the P-Values representing the statistical significance of these correlation coefficients, between three selected principal components and two S&P transparency and disclosure rankings. The principal components selected for this analysis are wordcount, informativeness and forward-looking words, which are component 1, 4 and 7 respectively. The S&P awards a composite rank based on firm-level information and an AR rank based on annual report data only. These ranks represent the extent of transparency and good disclosure practices in the company. A high transparency and disclosure rank represents a company with higher transparency and better disclosure. Though the AR rank and the composite rank are correlated with each other, they are not correlated with the informativeness and the forward looking principal components. Neither of the S&P ranks is significantly correlated with the occurrence of a lawsuit. However both informative and forward-looking disclosure are negatively correlated with lawsuits. Wordcount is positively correlated with lawsuits. P-Values for the correlation coefficients are in italics below the coefficient.

	AR Rank	Composite rank	Suit	PC4 Informativeness	PC7 Forward looking	PC1 Wordcount
AR Rank	1	0.3761	0.0371	0.0680	0.0234	-0.0503
		<.0001	*0.477*	*0.1918*	*0.6533*	*0.3346*
CompositeRank	0.3761	1	0.0001	0.0539	0.0286	0.0720
	<.0001		*0.998*	*0.3007*	*0.5834*	*0.1672*
Suit	0.0371	0.0001	1	-0.1088	0.0411	0.0494
	0.477	*0.998*		*<.0001*	*0.0029*	*0.0003*
PC4 Informativeness	0.0680	0.0539	-0.1088	1	-0.0723	0.0147
	0.1918	*0.3007*	*<.0001*		*<.0001*	*0.2879*
PC7 Forward looking	0.0234	0.0286	0.0411	-0.0723	1	-0.0098
	0.6533	*0.5834*	*0.0029*	*<.0001*		*0.4752*
PC1 Wordcount	-0.0503	0.0719	0.0494	0.0146	-0.0098	1
	0.3346	*0.1672*	*0.0003*	*0.2879*	*0.4752*	

Table 5.11

Comparison of informativeness component with S&P Transparency and Disclosure rank

Results from a horse race between the Principal Component 4 or Informativeness, and the S&P Transparency and Disclosure ranks are reported here. The S&P Transparency and Disclosure rank comes with an overall assessment called the Composite rank and an assessment based on disclosure in annual reports, called the AR rank. This table compares the ability of the S&P ranks and the principal components to predict litigation, and hence the dependent variable here is the occurrence of a lawsuit. This is a binary-valued discrete choice variable. Controls include other common determinants of lawsuits, and monitoring variables such as the existence of a blockholder and the proportion of independent directors on the board. Components 1 through 9 are the principal components based on content analysis of word categories. For a description of these principal components or their constituent word categories, please see Tables 5.7 and 5.8.

Dep. Var: Suit	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Intercept	-14.0047	<.0001	-15.941	0.0002	-11.1675	<.0001
Ln(Size)	0.3477	0.0018	0.4027	0.0011	0.3735	<.0001
Ln(ROE)	-0.1127	0.7139	2.0991	<.0001	0.9898	0.0732
Restatement	1.946	<.0001	-0.117	0.7068	-1.1579	<.0001
Fisc. Year returns	-1.5096	<.0001	-1.5703	<.0001	-0.3542	0.0256
Blockholder	1.2953	0.0448	1.8989	0.073	0.1294	0.5082
PC Indep. Dirs	1.3796	0.0762	1.3318	0.0831	1.9277	<.0001
Composite Rank	0.0813	0.6992	0.1008	0.6427		
AR Rank					0.0617	0.3373
Component1	-0.0139	0.7056			0.0614	0.007
Component2					0.1996	<.0001
Component3					-0.0223	0.5547
Component4	-0.1809	0.0113	-0.1907	0.0117	-0.0711	0.0677
Component5					0.1801	0.0007
Component6					-0.027	0.6576
Component7			-0.1665	0.2281	-0.0711	0.2505
Component8					-0.1998	0.0243
Component9					-0.1998	0.0212

Table 6.1

Determinants of informativeness: Governance structures

The table reports the results from a regression in which the dependent variable is the value of the informativeness principal component for a given 10-K document of a company. The value of informativeness is determined from a principal component analysis of the loadings of a 10-K document among various word categories. The independent variables are financial and monitoring characteristics of the company. The G-Index of antitakeover defenses and the E-Index or entrenchment index are used as independent variables in the regression. Compensation variables such as the total pay and the incentive-based compensation of the CEO are also included. Financial characteristics include size, current profitability, lagged and future profitability and lagged stock returns. The document length is also included as it may be related to informativeness of the document as well. The regression includes controls for year and industry effects and standard errors are adjusted for firm-level clustering.

	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Intercept	0.1417	0.7587	0.2094	0.6506	0.5752	0.2116	-0.0992	0.8329	0.2233	0.6559	1.1386	0.0383
Restatement	0.1556	0.4387	0.1687	0.404	0.3585	0.1	0.2125	0.4142	0.5009	0.0637	0.4166	0.5749
Ln(Marketcap)	-0.0182	0.4299	-0.0155	0.5052	-0.0411	0.0816	0.0125	0.5852	-0.0298	0.256	-0.0375	0.1912
Ln(ROE)	-0.6252	<.0001	-0.6444	<.0001	-0.5515	<.0001	-0.642	<.0001	-0.5074	<.0001	-0.9426	<.0001
Fwd LnROE	0.0183	0.7657	0.0198	0.7427	0.0669	0.2591	-0.0676	0.3181	-0.0262	0.6987	0.1297	0.1498
Fisc yr return	-0.1182	0.0016	-0.1172	0.002	-0.1073	0.0052	-0.1553	<.0001	-0.1388	0.0016	-0.13	0.0449
Lag LnROE	-0.5669	<.0001	-0.5653	<.0001	-0.4752	<.0001	-0.6229	<.0001	-0.524	<.0001	-0.3339	0.0037
Disc. Accruals					-0.0001	0.0005			0	0.0078	-0.0001	0.1784
Document length(PC1)			-0.0279	0.0006	-0.0063	0.4433	-0.0272	0.0018	-0.0103	0.2438	-0.0092	0.387
PC Indep Dirs	0.3371	0.0529	0.4247	0.0177	0.3829	0.0301			0.5179	0.0077	0.3831	0.0833
Total Dirs.	0.0143	0.202	0.0188	0.1028	0.0337	0.0149			0.0315	0.0283	0.046	0.0053
G-Index			-0.0249	0.0535	-0.0207	0.0852			-0.0169	0.1906		
E-Index											-0.0679	0.0208
Total pay CEO							1.825	0.1533	1.7994	0.1721		
Incentive pay CEO							-0.0353	0.6938	0.0353	0.7022		
Year dummies		Yes		Yes		Yes		Yes		Yes		Yes
FF industry dummies		Yes		Yes		Yes		Yes		Yes		Yes
Cluster adjusted SEs		Yes		Yes		Yes		Yes		Yes		Yes

Table 6.2

Monitoring structures as determinants of informativeness: Two stage least squares regression

This table shows the results of a Two-Stage Least Squares regression in which membership in the S&P index is used as an instrument for institutional monitoring. This instrument is used to do logistic regressions of the occurrence of a lawsuit on the institutional monitoring variable. First and second stage regression results are given in Panel A and B respectively. The three first stage regressions have as a dependent variable the proportion of shares held by institutions, the existence of a blockholder, and the number of block holders respectively. In all these three regressions, the dependent variable is S&P index membership and the coefficients in the table are the coefficients on this variable. P-Values and the adjusted R-squared are also displayed. In the second stage regressions, the informativeness measure is the dependent variable and the various independent variables are listed in the table.

First stage regression estimates			
Indep. Var.	Dep. Var.	Coefficient	P-Value
S&P Index	Prop. Institutional holdings	0.0368	0.0159
S&P Index	Blockholder	0.1537	<.0001
S&P Index	# Block holders	0.9864	<.0001

	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Intercept	-0.4272	0.0004	-0.4986	<.0001	-0.6015	<.0001	-1.3000	<.0001
Restatement	0.0550	0.6022	0.0010	0.9922	0.0002	0.9436	0.0060	0.0944
Ln(Marketcap)	-0.0239	0.0003	-0.0323	<.0001	-0.0136	0.043	0.0233	0.0045
Ln(ROE)	-0.1385	<.0001	-0.1708	<.0001	-0.3668	<.0001	-0.4183	<.0001
Future LnROE					0.0162	0.6254	0.0525	0.1982
Lag LnROE					-0.1531	<.0001	-0.1701	<.0001
Institutional holdings	0.5476	<.0001						
Blockholder			0.5865	0.0002				
# Blockholders					0.0615	0.0316		
Analyst coverage							0.0279	0.0056
Year dummies		Yes		Yes		Yes		Yes
FF industry dummies		Yes		Yes		Yes		Yes

References

Agrawal, Anup and Sahiba Chadha, 2005, Corporate governance and accounting scandals, *Journal of Law and Economics* 48, 371–406.

Bajaj, Mukesh, Sumon C. Mazumdar and Atulya Sarin, 2000, Securities class action settlements: An empirical analysis, Working paper, Securities Class Action Clearinghouse, Stanford University.

Bebchuk, Lucian, Alma Cohen and Allen Ferrell, 2004, Harvard Law School John M. Olin Center Discussion Paper No. 491.

Beneish, Messod, 1999, Incentives and penalties related to earnings overstatements that violate GAAP, *The Accounting Review* 74, 425-457.

Botosan, C., 1997, Disclosure level and the cost of equity capital, *The Accounting Review* 72, 323–350.

Brown, Stephen, Mark Finn and Stephen Hillegeist, 2000, Disclosure policies and the probability of informed trade, Research-in-progress, Northwestern University.

Bushee, Brian J. and Christopher F. Noe, 2000, Corporate disclosure practices, institutional investors and stock return volatility, *Journal of Accounting Research* 38, 171-202.

Carleton, Willard T., Michael S. Weisbach and Elliott J. Weiss, 1996, Securities class action lawsuits: A descriptive study, *Arizona Law Review* 491.

Chamberlain, Gary, 1980, Analysis of covariance with qualitative data, *Review of Economic Studies* 47, 225–238.

Cheng, Agnes, Denton Collins, and Henry Huang, 2003, The effect of the S&P T&D rankings on market beta, abnormal returns and earnings response coefficients in the period surrounding the report release date, Working paper, University of Houston.

Choi, Stephen J., 2005, Do the merits matter less after the private securities litigation reform act?, NYU Law & Economics Paper No. 03-04.

Clyde, Paul S., 1997, Do Institutional Shareholders Police Management?, *Managerial and Decision Economics* 18, 1-10.

Core, J. E., 2001, A review of the empirical disclosure literature, *Journal of Accounting and Economics* 31, 441-456.

Cox, C. T., 1985, Further evidence on the representativeness of management earnings. forecasts, *The Accounting Review* 60, 692-701.

Cox, James D. and Randall S. Thomas, 2003, SEC enforcement actions for financial fraud and private litigation: An empirical inquiry, Vanderbilt Law and Economics Research Paper No. 03-08.

Dechow, Patricia M., Richard G. Sloan and Amy P. Sweeney, 1995, Detecting earnings management, *The Accounting Review* 70, 193-226.

Dechow, Patricia M., Richard G. Sloan and Amy P. Sweeney, 1996, Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC, *Contemporary Accounting Research* 13, 1-36.

Demsetz, Harold, 1983, The structure of ownership and the theory of the firm, *Journal of Law and Economics* 26, 375-390.

Denis, David J., Paul Hanouna and Atulya Sarin, 2005, Is there a dark side to incentive compensation?, SSRN Working Paper.

Desai, H., C. Hogan and M. Wilkins, 2006, The reputational penalty for aggressive accounting: Earnings restatements and managing turnover, *The Accounting Review* 81, 83-112.

Easley, D., Soeren Hvidkjaer and Maureen O'Hara, 2002, Is information risk a determinant of asset returns?, *Journal of Finance* 57, 2185-2221.

Feroz, E.H., K. Park, and V.S. Pastena, 1991, The financial and market effects of the SEC's accounting and auditing enforcement releases, *Journal of Accounting Research* 2, 107-142.

Fich, Eliezer and Anil Shivdasani, 2004, The impact of stock option compensation for outside directors on firm value, *Journal of Business* 78, 2229–2254.

Field, Laura, Michelle Lowry, and Susan Shu, 2005, Does disclosure deter or trigger litigation?, *Journal of Accounting and Economics* 39, 487-507.

Fox, M. B., 1999, Required disclosure and corporate governance, *Law and Contemporary Problems* 62, 113-128.

Francis, Jennifer, Donna Philbrick and Katherine Schipper, 1994, Determinants and outcomes in class action securities litigation, Working paper, University of Chicago.

Francis, Jennifer, Dhananjay Nanda and Per Olsson, 2005, Voluntary disclosure, information quality, and costs of capital, Working paper, Duke University.

Gande, Amar and Craig M. Lewis, 2006, Shareholder initiated class action lawsuits: Shareholder wealth effects and industry spillovers, SSRN Working paper.

General Accounting Office, 2002, Financial statement restatements: Trends, market impacts, regulatory responses, and remaining challenges, GAO report 03-138.

Gillan, Stuart and Laura T. Starks, 2000, Corporate governance proposals and shareholder activism: The role of institutional investors, *Journal of Financial Economics* 57, 275-303.

Hartzell, Jay C. and Laura T. Starks, 2003, Institutional investors and executive compensation, *Journal of Finance* 58, 2351-2375.

Healy, Paul and Krishna Palepu, 1993, The effect of firms' financial disclosure strategies on stock prices, *Accounting Horizons* 7, 1–11.

Healy, Paul and Krishna Palepu, 2001, A review of the empirical disclosure literature, *Journal of Accounting & Economics* 31, 405–440.

Imhoff, Eugene A., Jr., 1978, The representativeness of management earnings forecasts, *Accounting Review* 53, 836-850.

Johnson, Keith L., 1997, Deterrence of corporate fraud through securities litigation: The role of institutional investors, *Law and Contemporary Problems* 60, 155-166.

Johnson, Shane A., Harley E. Ryan and Yisong S. Tian, 2003, Executive compensation and corporate fraud, Working paper, Louisiana State University.

Jones, Christopher L. and Seth E. Weingram, 1996, The determinants of 10B-5 litigation risk, Stanford Law School John M. Olin Program in Law and Economics Working paper no. 139.

Kasznik, Ron, 1999, On the association between voluntary disclosure and earnings management, *Journal of Accounting Research* 37, 57-81.

Kasznik, Ron and Baruch Lev, 1995, To warn or not to warn: Management disclosure and earnings management, *The Accounting Review* 70, 113-134.

Ke, Bin, 2003, The influence of equity based compensation on the CEO's incentive to report strings of consecutive earnings increases, Working paper, Duke University.

Kedia, Simi, 2003, Do executive stock options generate incentives for earnings management? Evidence from accounting restatements, SSRN Working paper.

Knutson P., 1992, Financial reporting into the 1990s and beyond, New York: Association for Investment Management and Research.

Lang, Mark and Russell Lundholm, 1993, Cross-sectional determinants of analyst ratings of corporate disclosures, *Journal of Accounting Research* 31, 246–271.

Lang, Mark and Russell Lundholm, 1996, Corporate disclosure policy and analyst behavior, *The Accounting Review* 71, 467–493.

Lev, Baruch, 1992, Information Disclosure Strategy, *California Management Review* 34, 9-32.

Livingston, James, 1996, Management-borne costs of fraudulent and misleading financial reporting, SSRN Working paper.

Maug, Ernst, 1998, Large shareholders as monitors: Is there a trade-off between liquidity and control?, *Journal of Finance* 53, 65-98.

Nagar, Venky, 1999, The role of the manager's human capital in discretionary disclosure, *Journal of Accounting Research* 37, 167-181.

Niehaus, Greg and Greg Roth, 1999, Insider trading, equity issues and CEO turnover in firms subject to securities class actions, *Financial Management* 22, 52-72.

Palmrose, Zoe-Vonna and Susan Scholz, 2004, The accounting causes and legal consequences of non-GAAP reporting: Evidence from restatements, *Contemporary Accounting Research* 21, 139-180.

Palmrose, Zoe-Vonna, V. J. Richardson, and S. Scholz, 2004, Determinants of market reactions to restatement announcements, *Journal of Accounting & Economics* 37, 59-89.

Peng, Lin and Ailsa Roell, 2004, Executive pay, earnings manipulation and shareholder litigation, SSRN Working Paper.

Richardson, Scott, Irem Tuna and Min Wu, 2002, Predicting earnings management: The case of earnings restatements, Working paper, University of Pennsylvania.

Shleifer, Andrei and Robert Vishny, 1986, Large shareholders and corporate control, *Journal of Political Economy* 94, 461-488.

Shleifer, Andrei and Robert Vishny, 1997, A survey of corporate governance, *Journal of Finance* 52, 737–783.

Skinner, Douglas J., 1994, Why firms voluntarily disclose bad news, *Journal of Accounting Research* 32, 38-60.

Skinner, Douglas J., 1997, Earnings disclosures and stockholder lawsuits, *Journal of Accounting and Economics* 23, 249-282.

Smith, Michael P., 1996, Shareholder activism by institutional investors: Evidence from CalPERS, *Journal of Finance* 51, 227-252.

Strahan, Philip, 1998, Securities class actions, corporate governance and managerial agency problems, Research paper 9816, Federal Reserve Bank of New York.

Summers, S. and J. Sweeney, 1998, Fraudulently misstated financial statements and insider trading: An empirical analysis, *The Accounting Review* 73, 131-146.

Tasker, Sarah, 1998, Bridging the information gap: Quarterly conference calls as a medium for voluntary disclosure, *Review of Accounting Studies* 3, 137-167.

Waymire, G., 1985, Earnings volatility and voluntary management forecast disclosure. *Journal of Accounting Research* 23, 268-295.

Weiss, Elliott J. and John S. Beckerman, 1995, Let the money do the monitoring: How institutional investors can reduce agency costs in securities class actions, *Yale Law Journal* 104, 2053-2127.

Wright, David W., 1996, Evidence on the relation between corporate governance characteristics and the quality of financial reporting, SSRN Working paper.

Wu, Min, 2002, Earnings restatements: A capital market perspective, Working paper, New York University.